The planning of headways in urban public transit

A. Grosfeld-Nir and J.H. Bookbinder

Department of Management Sciences, University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1

A basic issue in the planning of urban public transport is the determination of headways or inter-dispatch times. During each season, i.e. distinct time-period whose demand characteristics are constant, the following tradeoff must be considered. Dispatching too many vehicles on a route causes high operating costs, while too few vehicles may result in unsatisfactory levels of service. An appropriate policy on headways will help to balance resources between lines (routes) in peak-demand hours and will influence the total number of buses acquired by a transit company. Previous practice in industry usually bases the planning of headways upon satisfying service criteria on a "most-congested segment". This approach reduces the problem from that of studying a route to that of a single segment (stop), but thereby fails to account for other important information about the line's characteristics. In this article, we develop two new service criteria which consider the line as a whole: (1) "crowding-over-distance" takes into account discomfort resulting from a vehicle carrying too many passengers, and the corresponding distance travelled; and (2) "probability-of-failure", the frequency with which a waiting passenger fails to board due to lack of space. COD will be analyzed using simulation. POF will be related to a time-dependent Markov chain that is "inhomogeneous" in terms of distance along the route. Optimal headways are those which dispatch the smallest number of buses while meeting the particular service criterion. Models based on each of the two criteria are illustrated and applied to a number of routes of the Israeli transit company, DAN.

1. Introduction

It goes without saying that the planning of headways is important in urban public transportation. Transit management wishes to deliver a promised standard of service without incurring undue costs: if headways are too small (too many buses are dispatched), the company suffers due to excessive operating cost. Also, during peak hours, if more buses than necessary are assigned to one line, these buses will

*The preparation of this paper was partially supported by the Natural Sciences and Engineering Research Council of Canada, Grant No. OOP 05292.

© J.C. Baltzer AG, Science Publishers
be missing on another line, for usually, in peak times practically all buses are in operation. If, on the other hand, headways are too large, some service criteria may not be met, resulting in dissatisfied passengers who may choose alternative means of transportation.

Typical standards of service which we consider are related to crowding, i.e. the number of passengers on arriving vehicles, and the frequency that the bus is not too heavily loaded. A transit company will often plan headways by considering only a single critical (most-congested) stop or line segment. The logic behind this approach is that, by satisfying service standards for a most-congested segment, there is high probability that these standards will be satisfied for the entire line. This then reduces the problem of planning headways for a route, to the problem of planning headways for a single segment. The simplification is not without cost: a route with only one critical stop should not be applied the same dispatch policy as a route with two (or more) critical stops; nor should two lines whose single critical stop has similar characteristics necessarily have the same headway if the lines differ notably at their other stops.

Certainly a number of academic references offer service criteria involving more than a single critical stop. For example, Woodhull et al. [17] work separately with $S$ (the proportion of people in the bus that are standing) and $L$ (the load factor, i.e. number of people in bus ÷ number of seats). They develop service criteria in terms of relationships between $S$ and $L$. Ceder [4] advises use of the criterion, $\Sigma$(distance times number of passengers on bus). We remark that this is quite similar to our distance-based criterion (COD), except that only those route segments with ridership above a certain threshold contribute to COD.

The present study thus develops two models or criteria to determine headways on a route for each “season” (distinct demand-period of the day). The first criterion, which we term COD (crowding-over-distance, section 3), measures the discomfort to passengers when there is excessive crowding, in light of how long the crowding persists. The second criterion, POF (probability-of-failure, section 4), is based upon the likelihood that some users waiting at a particular stop will be unable to board a vehicle that is already heavily loaded. Both criteria take into account characteristics of the whole route and are generalizations of similar criteria for planning headways using a most-congested segment, as practiced at many transit authorities.

The models of this paper were in response to particular situations faced by the Israeli transit company, DAN. In 1990, DAN operated more than 1400 buses for up to 20 hours per day on 110 routes in Dan County which includes several cities, in particular Tel Aviv. Our two models were calibrated, tested, and applied to 30 of the most congested routes on the DAN system, in 79 seasons, including 40 peak seasons. We remark that throughout this article, the terms “route” and “line” will be used interchangeably.

Before proceeding to our models, we briefly discuss the relevant literature. Kocur and Henderson [10] obtain closed-form solutions for the optimal headway,
and the optimal route spacing and fare, for several objective functions. Kuah and Perl [11] found the optimal bus-stop spacing by minimizing total cost (i.e. of operator and user), but for fixed demand. Neither of those references considered time-dependent demand, which is quite important in our work. Newell [12] treated demand that was time dependent, but constant within each period (we will assume the same). For a fleet size that is given and fixed, Newell minimized user waiting time, subject to a constraint on transit vehicle capacity. Chang and Schonfeld [6] also studied time-dependent demand (and in fact, transit vehicle speeds that varied over time). They obtained closed-form solutions for analytic models which optimize the bus headways within each period.

Adamski [1] considers the service (boarding and alighting) provided to transit users at bus stops. His probabilistic models concentrate on those stochastic processes, the various probability density functions and/or the first two moments. The number of passengers on the bus is not at all discussed.

Adebisi [2] assumes that the number of passengers served at any stop is a stationary Poisson process. He finds headway variability is mainly due to traffic along the route and to bus-loading conditions. That paper, however, does not consider the variance in headways due to increases in ridership as the rush hour nears. Stephanedee et al. [15] consider transient (as well as steady-state) solutions in which they maximize the load factor, subject to constraints on the revenue/cost ratio, transit fare, and passenger waiting time. We remark that the first two constraints are not at all issues in our work. The final constraint, although closer to the present paper, is addressed in our study via the frequency of service, i.e. headway, which is our decision variable.

Tam and Seneviratne [16] employ a simulation model to study the effects of different headways. The other references cited to this point were analytic in nature. As noted by Powell and Sheffi [14], the results of a simulation model will be one or more confidence intervals for the output variable(s) of interest, whereas analytic models give more specific results that are limited by the particular assumptions made.

The numerical approach of Powell and Sheffi (see also Furth and Wilson [9]) is intermediate between the simulation and analytic models. In particular, Powell and Sheffi numerically obtain the distributions for each successive bus of its time of arrival and departure at the various stops on the route. The recursive procedure of those authors permits calculation, for example, of the distribution of passenger loads along the bus route. That distribution is a major concern of the present paper, for it enters into both of our service criteria.

2. Preliminaries

We begin by defining a route as consisting of $N$ segments and $N + 1$ stops; the $k$th segment, $k = 1, \ldots, N$, follows the $k$th stop (see figure 1). Typically, for
DAN, the number of stops per route is between 30 and 50 and the length of the route is about 12 km.

A season for a particular route is a time interval of the day, during which it is sensible to schedule constant headways. A season is thus a distinct demand-period. Each route may have six to ten of them, e.g. pre-AM rush; AM rush; mid-morning; noon break; mid-afternoon; PM rush; early evening; late evening. Our approach in this paper is that inter-departure times can be planned on a season-by-season basis. For a bus route during the given season, the rate of passenger arrivals will be constant (independent of time), but of course may vary by location along the route. Within that season, we will model the arrivals for boarding at each stop \( k \) on this line as a Poisson process with parameter \( \lambda_k \) (Chapman et al. [7]). Similarly, we assume that \( p_k \), the probability that an on-board passenger will depart at stop \( k \), is constant for that route and season.

For technical reasons, we will treat vehicle capacity as unlimited. This amounts to assuming that a renewal occurs at each stop right after a bus leaves there (i.e. none of the passengers waiting for a given bus were unable to board the previous vehicle). Thus, the probability that \( x \) passengers wish to board at that stop is the same for the next bus as it was for the bus that just left. Also, since occasionally we will recommend changes to headways, we implicitly assume that small differences in headways do not alter demand or alighting characteristics. For a fixed route and a given headway of \( t \), we denote:

- \( X_k \) = number of passengers waiting to board at stop \( k \), when a bus arrives;
- \( Y_k \) = number of passengers alighting at that stop; and
- \( N_k \) = number of passengers on board in segment \( k \).

Because the decision to alight at any stop is made independently by each passenger, it follows (Andersson and Scalia-Tomba [3], Powell and Sheffi [14]) that the conditional distribution of departures at stop \( k \), given the number of passengers on board just before this stop, is binomial. More precisely, the conditional distribution of \( Y_k \), given \( N_{k-1} \), is binomial with parameters \( (N_{k-1}, p_k) \). (In theorem 1, we prove that the unconditional distribution of \( Y_k \) is Poisson.)

The random variables \( X_k, Y_k \) and \( N_k \) will be important in section 3 when we study by simulation the consequences of the COD criterion. In section 4, we will consider POF by showing that the stochastic process corresponding to the number
of passengers on board, in consecutive route segments, is a non-homogeneous Markov chain. POF is the probability of first passage to a particular state related to the practical capacity of the bus. Numerical examples from DAN are given throughout the remainder of this paper. Section 5 presents our conclusions and suggestions for further research.

3. Crowding-over-distance

In this section, we develop the crowding-over-distance (COD) criterion. $N_k$ (as before) is the number of on-board passengers in segment $k$, which has length $L_k$. COD in segment $k$ is defined as

$$COD(k) = L_k \max\{0, N_k - H\},$$

where $H$ is the maximum number of passengers which travel with comfort, i.e. without excessive crowding. ($H$ is determined by transit Management through surveys of system users.) We define COD, for one particular trip, as

$$COD = \sum_k COD(k).$$

In our study, after consulting with experts and the client, we chose $H = 50$. (This is not inconsistent with the “optimal vehicle size” of 60 or 65 persons, determined analytically by Oldfield and Bly [13] for a public transit service.)

We denote by $t$ the headway and by $t^*$ the optimal headway. $t^*$ will minimize the number of buses assigned to a route, while satisfying a service constraint based upon COD.

In order to calibrate the COD criterion we chose several “control routes”: Routes numbered 4, 24, 46 and 48 in the DAN scheme were, in the expert opinion of Management, “optimal” at that time. By simulation we found the COD for various headways, including the one currently scheduled. As an example, table 1 presents the results obtained by simulating 1000 trips on Route 46 for each of three different headways. We observe here (and for the other control lines) that the frequency for which COD = 0 is at least 25%, and that the frequency of [COD = m] is decreasing in $m$. From these observations, we propose, with $g(t, m)$ as the empirical frequency of buses with COD = $m$, that

$$t^* = \sup\{t: g(t, m) \text{ is decreasing in } m \text{ and } g(t, 0) \geq 0.25\}. \quad (1)$$

That is, we require $t^*$ to be the largest headway so that

(a) The probability of COD = 0 is at least 25%.

(b) The probability of having COD = $m$ be decreasing in $m$. (Therefore, zero is the mode of COD.)
Table 1

The frequency for which COD = m for several headways on Route 46. For example, when headways are 7 minutes, the probability of COD between 10 and 15 is 0.013 (13 out of 1000 simulated bus runs). Note that the "present headway" is the average inter-departure time as measured on the street.

<table>
<thead>
<tr>
<th>COD = m</th>
<th>t = 7</th>
<th>t = 8.97</th>
<th>t = 9.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>828</td>
<td>304</td>
<td>96</td>
</tr>
<tr>
<td>0–5</td>
<td>116</td>
<td>249</td>
<td>130</td>
</tr>
<tr>
<td>5–10</td>
<td>38</td>
<td>156</td>
<td>135</td>
</tr>
<tr>
<td>10–15</td>
<td>13</td>
<td>109</td>
<td>140</td>
</tr>
<tr>
<td>15–20</td>
<td>3</td>
<td>72</td>
<td>131</td>
</tr>
<tr>
<td>20–25</td>
<td>2</td>
<td>44</td>
<td>95</td>
</tr>
<tr>
<td>25–30</td>
<td>0</td>
<td>35</td>
<td>87</td>
</tr>
<tr>
<td>30–35</td>
<td>0</td>
<td>12</td>
<td>57</td>
</tr>
<tr>
<td>35–40</td>
<td>0</td>
<td>10</td>
<td>53</td>
</tr>
<tr>
<td>40–45</td>
<td>0</td>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td>above 45</td>
<td>0</td>
<td>5</td>
<td>43</td>
</tr>
</tbody>
</table>

3.1. AN EXAMPLE

As another example, consider Line 66 (inbound direction) during the peak morning hour, 6:45–8:30 a.m. This line has 37 stops and the present headways are 3.5 minutes. The data (table 2) consists of the lengths of the route segments, the rates of arrival at each stop, and the alighting probabilities there.

It is not difficult to simulate a bus run and thus to calculate the resulting COD. First, find the number $X_1$, of passengers boarding at stop 1, from a Poisson distribution with parameter $\lambda_1 t$ (t = headway) and set $N_1 = X_1$ (recall that $N_k$ represents the number of on-board passengers in segment $k$). Then draw the number $Y_2$, of passengers alighting at stop 2, from a binomial distribution with parameters ($N_1, p_2$). Now find the number $X_2$, of passengers boarding at stop 2, from a Poisson distribution with parameter $\lambda_2 t$ and set $N_2 = N_1 + X_2 - Y_2$, etc. Finally, obtain COD($k$) = $L_k \max\{0, N_k - H\}$, and COD = $\Sigma_k$COD($k$). The results are summarized in table 3.

From table 3, it is clear that COD = 0 occurs more than 25% of the time, and that the frequency $f$ is decreasing in the COD. Thus, the headway need not be reduced. To check whether it is desirable to increase headways, we simulated the same route with $t = 4$ minutes.

From table 4, we conclude that this increased headway is not optimal, because the frequency first increases in COD before beginning to decrease; and also because the frequency for which COD = 0 is less than 25%. The currently-scheduled headway of 3.5 minutes is therefore recommended.
The data for Line 66. For example, the arrival rate at stop 8 is $\lambda_8 = 0.61$ passengers per minute; alighting probability at this stop is $p_k = 0.11$; the distance to the next stop (length of segment number eight) is $L_8 = 580$ meters. The overall length of the route is 16.94 km.

<table>
<thead>
<tr>
<th>Stop</th>
<th>$L_k$</th>
<th>$\lambda_k$</th>
<th>$p_k$</th>
<th>Stop</th>
<th>$L_k$</th>
<th>$\lambda_k$</th>
<th>$p_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>210</td>
<td>4.65</td>
<td>0.00</td>
<td>20</td>
<td>370</td>
<td>1.10</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>4.59</td>
<td>0.00</td>
<td>21</td>
<td>380</td>
<td>0.63</td>
<td>0.05</td>
</tr>
<tr>
<td>3</td>
<td>330</td>
<td>1.37</td>
<td>0.00</td>
<td>22</td>
<td>1000</td>
<td>1.44</td>
<td>0.11</td>
</tr>
<tr>
<td>4</td>
<td>380</td>
<td>2.83</td>
<td>0.02</td>
<td>23</td>
<td>460</td>
<td>0.57</td>
<td>0.15</td>
</tr>
<tr>
<td>5</td>
<td>490</td>
<td>1.92</td>
<td>0.02</td>
<td>24</td>
<td>540</td>
<td>0.44</td>
<td>0.11</td>
</tr>
<tr>
<td>6</td>
<td>470</td>
<td>1.36</td>
<td>0.01</td>
<td>25</td>
<td>350</td>
<td>0.28</td>
<td>0.10</td>
</tr>
<tr>
<td>7</td>
<td>450</td>
<td>0.90</td>
<td>0.02</td>
<td>26</td>
<td>280</td>
<td>0.30</td>
<td>0.08</td>
</tr>
<tr>
<td>8</td>
<td>580</td>
<td>0.61</td>
<td>0.11</td>
<td>27</td>
<td>560</td>
<td>0.39</td>
<td>0.14</td>
</tr>
<tr>
<td>9</td>
<td>270</td>
<td>0.02</td>
<td>0.03</td>
<td>28</td>
<td>510</td>
<td>0.30</td>
<td>0.06</td>
</tr>
<tr>
<td>10</td>
<td>420</td>
<td>0.03</td>
<td>0.09</td>
<td>29</td>
<td>500</td>
<td>0.07</td>
<td>0.15</td>
</tr>
<tr>
<td>11</td>
<td>420</td>
<td>0.13</td>
<td>0.07</td>
<td>30</td>
<td>1480</td>
<td>0.09</td>
<td>0.21</td>
</tr>
<tr>
<td>12</td>
<td>240</td>
<td>0.19</td>
<td>0.08</td>
<td>31</td>
<td>300</td>
<td>0.02</td>
<td>0.24</td>
</tr>
<tr>
<td>13</td>
<td>630</td>
<td>0.76</td>
<td>0.08</td>
<td>32</td>
<td>300</td>
<td>0.00</td>
<td>0.24</td>
</tr>
<tr>
<td>14</td>
<td>320</td>
<td>2.55</td>
<td>0.04</td>
<td>33</td>
<td>500</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>15</td>
<td>580</td>
<td>1.72</td>
<td>0.04</td>
<td>34</td>
<td>730</td>
<td>0.00</td>
<td>0.13</td>
</tr>
<tr>
<td>16</td>
<td>210</td>
<td>0.89</td>
<td>0.02</td>
<td>35</td>
<td>200</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>17</td>
<td>720</td>
<td>0.50</td>
<td>0.09</td>
<td>36</td>
<td>200</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>18</td>
<td>480</td>
<td>1.87</td>
<td>0.09</td>
<td>37</td>
<td>0</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>19</td>
<td>260</td>
<td>1.13</td>
<td>0.05</td>
<td></td>
<td></td>
<td>16,940</td>
<td></td>
</tr>
</tbody>
</table>

Table 3

The distribution of COD for Line 66 when $t = 3.5$ minutes (currently scheduled headway). For example, the frequency of COD between 5 and 10 is $f = 10.4\%$ (104 out of 1000 simulated bus runs).

<table>
<thead>
<tr>
<th>COD</th>
<th>0-5</th>
<th>5-10</th>
<th>10-15</th>
<th>15-20</th>
<th>20-25</th>
<th>25-30</th>
<th>30-35</th>
<th>35-40</th>
<th>40-45</th>
<th>45-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>482</td>
<td>246</td>
<td>104</td>
<td>55</td>
<td>46</td>
<td>20</td>
<td>14</td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4

The distribution of COD for Line 66 when $t = 4$ minutes.

<table>
<thead>
<tr>
<th>COD</th>
<th>0-5</th>
<th>5-10</th>
<th>10-15</th>
<th>15-20</th>
<th>20-25</th>
<th>25-30</th>
<th>30-35</th>
<th>35-40</th>
<th>40-45</th>
<th>45-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>135</td>
<td>187</td>
<td>143</td>
<td>91</td>
<td>81</td>
<td>66</td>
<td>44</td>
<td>34</td>
<td>36</td>
<td>46</td>
</tr>
</tbody>
</table>
3.2. SUMMARY OF RESULTS

To further explore the COD criterion, we examined 23 different routes, mostly for peak-demand hours. The results are summarized in table 5.

We chose to recommend a change in headways only if the difference between the current headway and the optimal one was at least 10%. Out of the 23 lines of table 5:

(a) Nine were recommended for a decrease in headways (more buses).
(b) Five were recommended to have an increase in headways (fewer buses).
(c) Nine were suggested to remain unchanged.
To gain additional insight, we tested the performance of COD against the most-congested-segment criterion in 56 seasons on Routes 1, 23, 46, 61, and 67. (It was for these lines that we had the most recent data.) Our results indicate that 29 seasons required greater headways (less service); 14 merited a decrease in headways; and 13 seasons should remain unchanged.

4. Probability-of-failure

In this section, we consider the “probability-of-failure” (POF) criterion. Essentially, POF is the probability that a bus will fail to comfortably board all the passengers waiting at a stop, for at least one stop along the route. More precisely, let \( M - 1 \) be the maximum number of passengers that could be comfortably carried on a bus. We say that a “failure” occurred if the number of passengers on the bus, minus the number of passengers alighting at the next stop, plus the number of passengers waiting to board there is \( M \) or greater, for one or more stops along the route. Because we model the standing-capacity of a bus as unlimited, let us define the POF as the probability that the number of passengers on the bus is \( M \) or more, in at least one line segment.

To obtain POF, we start by showing that the stochastic process \( \{N_k, k = 1, \ldots, N\} \), i.e. the number of passengers in successive route segments, is a Markov chain. Here, the state space \( 0, 1, 2, \ldots \) is infinite, since we assume for \( N_k \) that there is unlimited capacity on the bus. Computation of the transition probabilities of \( N_k \) will aid us in working with the finite-state Markov chain \( \{N'_k, k = 1, \ldots, N\} \). Its state space is \( 0, 1, \ldots, M \), where \( M \) is an absorbing state. After that, use of “first passage” transition probabilities enables one to find POF.

Recalling the remarks at the end of section 2 on the random variables \( X_k \), \( Y_k \) and \( N_k \), we have

**THEOREM 1**

a) The (unconditional) number of passengers \( Y_k \) alighting at stop \( k \) is a Poisson process with rate \( \mu_k \), where \( \mu_1 = 0 \) (since \( p_1 = 0 \)) and

\[
\mu_{k+1} = [\lambda_1 + \lambda_2 + \ldots + \lambda_k] - (\mu_1 + \mu_2 + \ldots + \mu_k) \text{ if } k \geq 1.
\]

b) \( N_k \), the number of passengers continuing on the bus in segment \( k \), is a Poisson process with rate:

\[
(\lambda_1 + \lambda_2 + \ldots + \lambda_k) - (\mu_1 + \mu_2 + \ldots + \mu_k).
\]

**Proof**

The number of passengers in segment 1 (having arrived at stop 1) is, by hypothesis, a Poisson process with rate \( \lambda_1 = \lambda_1 - \mu_1 \). When the vehicle proceeds to
stop 2, the probability of any given passenger alighting is $p_2$ independent of the other passengers. Therefore, the number of passenger alighting is a Poisson process with rate $\mu_2 = \lambda_1 p_2$ and, similarly, the number of passengers remaining is a Poisson process with rate $\lambda_1 (1 - p_2) = \lambda_1 - \mu_2$ (see, for example, Çınlar [8, chapter 4]). Now, $N_2$, the number of passengers boarding there is again, by hypothesis, Poisson with rate $\lambda_2$, independent of the number already on board. Therefore, the number of passengers in segment 2 is a Poisson process with rate

$$\lambda_2 + [\lambda_1 (1 - p_2)] = \lambda_1 + \lambda_2 - (\mu_1 + \mu_2)$$

(since $\mu_1 = 0$). The rest follows by induction.

\[ \square \]

**Theorem 2**

The stochastic process $\{N_k; k = 1, 2, \ldots, N\}$ is a Markov chain. Let $r_k(i, j)$ denote the transition probabilities

$$r_k(i, j) = P(N_k = j \mid N_{k-1} = i).$$

Then for a headway of $t$

$$r_k(i, j) = \sum_{y=0}^{i} \binom{i}{y} p_k^y (1 - p_k)^{i-y} e^{-\lambda_k t} \frac{(\lambda_k t)^{j-i+y}}{(j-i+y)!}. \tag{2}$$

**Proof**

First note that

$$N_k = N_{k-1} + X_k - Y_k.$$  

Observe that $X_k$ is independent of any of the $N_k$ and that $Y_k$ only depends on $N_{k-1}$ (and not on $N_{k-2}, N_{k-3}, \ldots$). Therefore, $N_k$ also depends only on $N_{k-1}$. Now, using the notation of section 2,

$$r_k(i, j) = P(N_k = j \mid N_{k-1} = i)$$

$$= \sum_{y} P(N_k = j \mid Y_k = y, N_{k-1} = i) P(Y_k = y \mid N_{k-1} = i)$$

$$= \sum_{y} P(X_k = j - i + y) P(Y_k = y \mid N_{k-1} = i).$$

Finally, to obtain the result recall that $X_k$ is Poisson with parameter $\lambda_k t$ and that the conditional distribution of $Y_k$, given $N_{k-1} = i$, is binomial with parameters $(i, p_k)$.

\[ \square \]

Now let $q_k(i, j)$ be the transition probabilities of the finite-state Markov chain $N'_k$, with state space $0, 1, \ldots, M$. State $M$ is absorbing, i.e. $q_k(M, M) = 1$, and
\[ q_k(i, j) = r_k(i, j) \quad \text{if} \quad i, j = 0, 1, \ldots, M - 1 \]

and

\[ q_k(i, M) = \sum_{j \geq M} r_k(i, j), \quad i = 0, 1, \ldots, M - 1. \]  

Note that the Markov chain \( N_k \) has the same transition matrix as the Markov chain \( N_k \) for columns 0, 1, \ldots, \( M - 1 \). However, column \( M \) of \( N_k \) consists of the probability that the number of passengers (according to \( N_k \)) is \( M \), at least.

With this notation, POF is thus the probability that the (finite-state) Markov chain will ever visit the absorbing state \( M \). To calculate POF, let \( F_k(M) \) be the first-passage probability, i.e., the probability that \( M \) is attained for the first time at step \( k \). Therefore,

\[ \text{POF} = F_1(M) + F_2(M) + \ldots + F_N(M), \]  

where we have taken account that the “time horizon” is finite and equals \( N \), the number of route segments. (\( N + 1 \) is the number of stops.)

Finally, let \( \pi_k(j), k = 1, \ldots, N, \) be the probability of having \( j \) passengers on board in segment \( k \). Note that, for a fixed headway of \( t \), \( \pi_1(j), j = 0, 1, \ldots, M, \) is simply the probability that \( j \) passengers arrive at stop 1 (determined from a Poisson distribution with parameter \( \lambda_1 \)). Also,

\[ \pi_k(j) = \sum_i \pi_{k-1}(i) q_k(i, j), \quad k = 2, \ldots, N. \]  

Because \( F_1(M) = \pi_1(M) \), the \( F_k(M) \) for \( k > 1 \) can be computed recursively via

\[ F_k(M) = \sum_{i=0}^{M-1} \pi_{k-1}(i) q_k(i, M). \]  

In summary, the calculation of POF requires the following steps (we fixed \( M = 81 \) throughout our numerical work):

(a) Find the transition probabilities \( r_k(i, j) \) using (2) and \( q_k(i, j) \) from (3).
(b) Determine \( \pi_k(j) \) recursively employing (5).
(c) Obtain \( F_k(M) \) as in (6) and the POF utilizing (4).

**Comment**

The probability of failure can be obtained somewhat more easily as

\[ \text{POF} = \pi_N(M). \]

(Simply note that the events “the Markov chain is in (the absorbing) state \( M \) after \( N \) stops” and “state \( M \) has been reached in stop \( N \) or before”, are equivalent.) While eliminating step (c), equation (7) is not much simpler because most of the computational burden lies in (b) (equation 5). We have employed step (c), i.e. equation
(4) rather than (7), for the following reason. The $F_k(M)$ from (6) furnish a measure of the probability of failure at a particular stop. Action may be taken once it is realized that $F_k(M)$ is large for a given stop $k$.

Now, in order to understand and interpret the values of POF, we computed it for the same control lines as in section 3. In addition, we analyzed several other routes using headways resulting from the COD criterion. POF in each of these cases turns out to be 5 out of 1000, or less. Consulting with transit Management and transport experts who viewed these results, it was concluded that optimal headways should be such that POF ≤ 0.005.

It may appear that this is too stringent a requirement on POF. In reality, many instances of overcrowded buses result from delays due to traffic or an accident. (Those sources of delay are not considered in this paper.) If buses had their own right-of-way, independent of vehicular traffic, it would be correct to interpret POF as an upper bound that a passenger will encounter an overcrowded bus. Under those conditions on right-of-way, the value 0.005 does not seem unreasonable.

4.1. AN EXAMPLE

As an example, consider Line 44, which has 22 stops. The present (observed) headway, during the morning peak hours (6:45–8:30), is $t = 6.57$ minutes. Arrival rates are

$$
\lambda_k = 1.07; \ 0.56; \ 0.51; \ 0.75; \ 0.74; \ 4.26; \ 2.18; \ 1.12; \ 0.47; \ 0.67; \ 0.40; \ 0.47; \ 0.63; \ 0.23; \ 0.32; \ 0.53; \ 0.17; \ 0.16; \ 0.07; \ 0.13; \ 0.02; \ 0.00.
$$

Alighting probabilities are:

$$
p_k = 0.00; \ 0.00; \ 0.00; \ 0.00; \ 0.00; \ 0.01; \ 0.00; \ 0.01; \ 0.01; \ 0.02; \ 0.07; \ 0.03; \ 0.09; \ 0.10; \ 0.05; \ 0.10; \ 0.01; \ 0.03; \ 0.04; \ 0.22; \ 0.33; \ 1.00.
$$

With these data and $M = 81$, steps (a)–(c) yield the results in table 6 for $F_k(M)$ and POF.

Table 6

<table>
<thead>
<tr>
<th>$k$</th>
<th>$F_k(M)$</th>
<th>$k$</th>
<th>$F_k(M)$</th>
<th>$k$</th>
<th>$F_k(M)$</th>
<th>$k$</th>
<th>$F_k(M)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0.0341</td>
<td>13</td>
<td>0</td>
<td>19</td>
<td>5E−7</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>8</td>
<td>0.1280</td>
<td>14</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>9</td>
<td>0.0842</td>
<td>15</td>
<td>4E−6</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>10</td>
<td>0.1207</td>
<td>16</td>
<td>3E−6</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>11</td>
<td>0.0103</td>
<td>17</td>
<td>7E−6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3E−5</td>
<td>12</td>
<td>0.0401</td>
<td>18</td>
<td>3E−6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

POF = 0.4147
Table 7
First passage (absorption) probabilities for Line 44 for headway of 5 minutes.

<table>
<thead>
<tr>
<th>k</th>
<th>$F_k(M)$</th>
<th>k</th>
<th>$F_k(M)$</th>
<th>k</th>
<th>$F_k(M)$</th>
<th>k</th>
<th>$F_k(M)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>7</td>
<td>8.0E-6</td>
<td>13</td>
<td>0</td>
<td>19</td>
<td>2.7E-10</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>8</td>
<td>238E-6</td>
<td>14</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>9</td>
<td>554E-6</td>
<td>15</td>
<td>1.7E-8</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>10</td>
<td>2067E-6</td>
<td>16</td>
<td>6.5E-9</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>11</td>
<td>1608E-6</td>
<td>17</td>
<td>6.2E-9</td>
<td></td>
<td>POF = 0.003706</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>12</td>
<td>679E-6</td>
<td>18</td>
<td>1.6E-8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6 reveals that POF is too high, i.e. headways are too large. Reducing the headway to 5 minutes we obtain the results of table 7. This headway was thus recommended. Incidentally, the same value $t = 5$ minutes was suggested by the COD criterion (see table 5). We shall now see that this need not always be the case.

4.2. RESULTS USING POF FOR SEVERAL LINES

**Line 4; season 6:45–8:30; current headway = 4.45 minutes.**

In this case, we obtain POF = 0, the same POF as obtained for a headway of 5 minutes. When the headway is 6 minutes, we find

$$F_{k=1,...,15} = 0; 0.00108; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0.$$ 

Therefore, POF = 0.00108. The recommended headway here is 6 minutes, different from that obtained by COD (5 minutes).

**Line 31; season 6:45–8:30; current headway = 5.65 minutes.**

We calculate

$$F_{k=1,...,17} = 0; 0; 0; 0; |32E-6; 2654E-6; 17119E-6; 3818E-6; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0.$$ 

for a POF of 0.024. Since this exceeds 0.5%, we tried headway = 5 minutes, and found

$$F_{k=1,...,17} = 0; ...; 0; 59E-6; 82E-6; 20E-6; 0; ...; 0.$$ 

Now the POF = 0.001. (The headway suggested under COD was 4.5 minutes.)
Line 57; season 6:45–8:30.

In this case, we used the headway of 19 minutes which was recommended by COD. The results are:

\[ F_{k=1,\ldots,22} = 0; \ldots; 0; 3E-6; 15E-6; 0.9E-6; 11E-6; 136E-6; 62E-6; 0; \ldots; 0, \]

for a POF = 0.000227.

Line 9; season 7:00–9:00.

For this route as well, we employed the same headway (4 minutes) suggested by the COD. The results are:

\[ F_{k=1,\ldots,30} = 0; \ldots; 0; 6.6E-6; 138E-6; 644E-6; 121E-6; 218E-6; 431E-6; 795E-6; 0; \ldots; 0. \]

The POF = 0.002354.

5. Concluding remarks

In this paper, we have suggested and analyzed two new service criteria for determining the appropriate headways in urban public transit. Crowding-over-distance (COD) weights the amount of passenger discomfort by the distance over which the crowding persists. Probability-of-failure (POF) is related to the likelihood that an arriving bus will have insufficient space for some users who wish to board; POF is modeled as a transition into a particular absorbing state of an inhomogeneous Markov chain.

The same input goes into determining the two criteria. Both COD and POF take into account characteristics of an entire route, rather than only a most-congested segment of that route. Greater satisfaction of transit users is indicated by smaller values of COD and smaller values of POF.

It is thus reasonable that COD and POF can often suggest similar headways; comparison of tables 5 and 7 shows that this is indeed the case for Route 44. The recommendations are also the same for Routes 9 and 57 (section 4.2). However (sections 3.2 and 4.2), the recommended headways differ slightly for Routes 4 and 31, for example. Variations between the two criteria will sometimes emerge for the following reason.

COD represents the disutility of a passenger presently on a vehicle; POF takes the point of view of a prospective passenger. Such a person is not yet on the bus and, for larger values of POF, s/he is less likely to be able to board comfortably. In fact, such occasional divergence may be desirable. It is well known that after a long wait, prospective passengers will try their best to board the first arriving vehicle, even though COD may be greatly reduced on the following vehicle, only
a few minutes behind. A rider already on a bus seems less bothered by a small increase in crowding, than someone waiting is bothered by a small increase in POF. It is thus helpful to have separate criteria representing the two points of view, in light of the political nature of most public-transit decisions.

We feel the positive results of this paper show that further research is merited on the preceding service criteria. Let us mention three possibilities for additional study. The first concerns the definition of COD(κ), which was given in section 3 as initially constant (zero), followed by a function linear in the number of passengers \( N_k \) on board in segment \( k \). It may be desirable to experiment instead with a convex increasing function that is zero only at the origin and differentiable everywhere. Secondly, it would also be interesting to consider a POF-service standard (maximum value for this probability) which depends on time-of-day (season).

Thirdly, either criterion of choice (COD or POF) may recommend a pattern of vehicle headways that could be difficult to schedule on the street. We argued in section 1 that not only should a single route be considered at a time, but also that headways could be chosen for one season at a time. It could happen that (a) the suggested headways vary greatly between adjacent seasons, or (b) there are too many seasons (so that headways as a function of time of day may have more than one increasing region, or more than one decreasing region). For either (a) or (b), union constraints on driver work assignments may permit scheduling the recommended sequence of headway-seasons, only with excessive costs of “penalties” or overtime. By restricting attention to the detailed union contract at a particular transit company, one could consider whether limiting the variation in season-headway patterns could actually aid in the scheduling of buses throughout those seasons. The net result would be some interesting tradeoffs between the cost (of driver work assignments) and service (COD or POF) to customers.

To conclude, we offer several remarks on the relationships between the two criteria and their respective parameters \( H \) and \( M \). Let \( t^* \) denote the optimal headway as recommended by COD. We have found that for our choice \( H = 50 \), a headway of \( t^* \) leads to POF always slightly smaller than 0.005. (For this \( H \), COD thus appears a bit more stringent than POF.) Also for headway \( t^* \), the average number of passengers in the most-congested segment is typically \( \approx 55 \), so (theorem 1) its standard deviation is \( \sqrt{55} \). A three-sigma upper limit for the number on the bus would thus approach 80 in the most-congested segment. This is satisfyingly close to the value of \( M \) we employed in studying POF, although we selected that value in a different way.

References


