HOW TO IMPROVE HIGH-FREQUENCY BUS SERVICE RELIABILITY THROUGH SCHEDULING

Ms. Grace Fattouche

Abstract
This paper outlines a scheduling process for improving high-frequency bus service reliability based on a model which uses Automatic Vehicle Location and Automatic Passenger Count data.

Developing a schedule for high-frequency bus routes involves balancing the costs to the passengers and the cost to the transit agency. Passengers are interested in short travel times and short, reliable waiting times. In order to assess the trade-off between trip speed and reliability, transit planners need to follow a clear scheduling process, which explicitly projects and evaluates the tradeoffs between overall travel time and reliability. The proposed model estimates the cost of any given schedule for waiting passengers, onboard passengers and the transit agency based on the existing variability in running times and headways. The scheduling process involves finding the time point schedule which minimizes the total cost with the help of the model. The model is based on two critical hypotheses: (i) consecutive bus vehicle trips are independent and (ii) consecutive segment running times for a particular bus trip are independent. These two critical hypotheses were shown to be true on two high-frequency bus routes analysed in Chicago; Route 95E and 85. For each route, the schedule which minimises the total passenger cost was determined. The schedules obtained with this generalised cost minimisation approach showed improved reliability and overall passenger service quality – at the same operating cost compared to the current schedules on both routes, as well as compared to traditional approaches. A sensitivity analysis has shown that in most cases the generalised cost minimisation schedule can significantly improve reliability and overall passenger service quality over traditional approaches.

INTRODUCTION
Bus service reliability is critical to passengers who are counting on their bus service to be on time. Depending on the type of bus route, low-frequency or high-frequency, passengers perceive reliability differently. On low-frequency bus routes (i.e., routes whose headway is more than ten-fifteen minutes), passengers usually arrive at stops based on a published schedule [1]. Consequently, a reliable service for the passengers on a low-frequency bus route is a service where buses arrive at stops on time (i.e., where the difference between the actual arrival time and the scheduled arrival time is very small). On low-frequency bus routes, consecutive trips are usually independent because of the long headways between runs and unreliability rarely propagates to following buses.

Conversely, on high-frequency bus routes, with headways of less than ten-fifteen minutes passengers usually arrive at stops randomly. On these bus routes, a reliable service is one where their expected waiting time is small or, in the ideal case, where the headway is constant and equal to the scheduled headway. On high-frequency bus routes, unreliability propagates easily to following buses because of the short headways and the interaction between buses. The consequences of unreliability for passengers include crowding, longer waiting time, missed appointments, higher travel time uncertainty and bus bunching. An unreliable service on a high-frequency bus route also affects the transit agency in terms of the number of vehicles required, overtime and increased operating costs for standby drivers, and in lost revenue due to reduced ridership [2].

Transit planners acknowledge that running time variability is an important cause of unreliability since it affects the overall on-time performance and the headway variability. This paper proposes a model-based approach for improving bus service reliability by adjusting schedules - one of the most common approaches among preventive strategies - in order to reduce running time variability and, consequently, reduce headway variability. A schedule
indicates to the operator the scheduled departure time from each time point\(^1\) along the route. Operators are instructed not to leave a time point early when a schedule-based holding strategy\(^2\) is enforced. The holding control strategy applies to each time point excluding the terminals where the operator may rest after all passengers get off the bus but has to leave if he arrives after the scheduled departure time.

Fortunately, with the development of technologies such as Automatic Vehicle Location (AVL) and Automatic Passenger Counters (APC), larger amounts of accurate data are now becoming available to evaluate and adjust schedules. Before the recent development of Automatic Data Collection (ADC) systems, and even today because of a lack of research in the area, transit planners use "rules of thumb" to develop "workable" schedules. These rules of thumb often vary from one transit agency to another. However, one common approach [3] is to: (1) Set running times at a level where at least 50% or 65% - depending on the transit agency - of trips would have sufficient time to complete a route or a segment, (2) Set layovers (or recovery time) at a level that would allow at least 90% to 95% of operators to depart their next trip on time.

**MODEL STRUCTURE**

Developing a schedule is quite complex since passengers are interested in a reliable service and in short in-vehicle travel time while the transit agency is also interested in minimizing operating costs. An intuitive approach to improve bus service reliability would be to allocate more scheduled running time between time points and more layover time at terminals. However, allocating more scheduled running time between time points increases passenger in-vehicle travel time to through passengers\(^3\) when schedule-based holding is enforced. Similarly, allocating more layover time at terminals increases operating costs if the transit agency wants to maintain the scheduled headway without increasing the number of buses in operation. Consequently, the proposed schedule design must balance improvements in service reliability against the increased cost both for the through passengers and for the transit agency. To evaluate the trade-off between reliability, speed and operating cost, a simple cost model is used to translate the reliability and the passenger service quality into economic values. Thus, the costs to waiting passengers, onboard passengers and to the agency can be found for each schedule tested.

Since it is not possible to experimentally test several schedules, we have to model how a new schedule is going to impact headway variability which is our measure of bus service reliability. Assuming that operators follow a schedule-based holding strategy, the model gives the actual arrival and departure times at each time point for each vehicle on a route with \(n\) time points, given a proposed time point schedule and the current scheduled headway. From the actual departure times from each time point obtained from the model, the actual headway and the headway variability can be computed. Thus, the service reliability for the passengers can be estimated. From the running times developed by the user and the number of buses to be operated on the route, the model derives the total scheduled layover time for the route. The amount of time allocated at each terminal will determine the headway variability and consequently the reliability experienced by passengers.

In summary, the descriptive model proposed here simulates the scheduled departure time from each time point along a route with a new schedule. It consequently estimates the passenger waiting time cost, the passenger in-vehicle cost and the operating cost of a proposed schedule. By systematically varying the schedule parameters, the schedule which minimises the passenger total cost can be identified. The inputs to the model are: the current AVL and APC data, the scheduled segment running times, the scheduled headway and the scheduled number of buses. The outputs are: the waiting passenger cost, the in-vehicle travel time cost and the operating cost.

\(1\) A time point is a bus stop at which the scheduled departure time is given to the operators.

\(2\) The schedule-based holding strategy is a self-monitoring measure in which the driver must hold his or her bus at a time point until its scheduled departure time if arriving early, and departs from the time point immediately upon completion of passenger processing if arriving late.

\(3\) Through passengers are the passengers who do not get off at the time point and consequently experience the holds at the time point.
The step-by-step process of the model is as follows:

1. The user inputs:
   - the segment running time distributions (from the departure time at the starting time point to the arrival time at the next time point) and the segment cumulative density functions based on the AVL data.
   - the scheduled segment running times by choosing on each segment the percentage of trips which can operate within the scheduled segment running time. If, for example, the user wants to evaluate a segment scheduled running time which is the 60th percentile of the segment cumulative density function, the scheduled running time on that segment should be 1.5 min, as shown in Figure 1.

   ![Figure 1. Segment Cumulative Density Function](image)

2. The model implements the scheduled running times and scheduled headway chosen by the user from the beginning of the day. It, then, simulates the vehicles' arrival and departure times from each time point on the route throughout the day up to the hour of operation studied, assuming that the same buses are used continuously.

3. For a given trip, the segment running time (from the departure time at the starting time point to the arrival time at the ending time point) is drawn randomly from the segment running time distribution.

4. The model calculates the departure times of each trip from each time point. With a schedule-based holding strategy enforced, the departure time from a time point is equal to the scheduled departure time if the vehicle arrived before the scheduled departure time or equal to the sum of the arrival time and dwell time at the time point if the vehicle arrived after the scheduled departure time.

5. From the arrival and departure times, the model calculates the average and coefficient of variation for arrival and departure headways at each time point.
6. Also, from the departure times, the model calculates the average time vehicles spend on each segment (from the departure time at the starting time point to the departure time at the next time point).

7. From the arrival and departure times at each terminal, the model calculates the average recovery time.

8. The passenger waiting time cost on each segment of the route is calculated by the model during the hour of operation studied. It is a function of the average headway, headway coefficient of variation and the expected number of passengers waiting on the segment weighted by the unit waiting time cost. (see [4], pg. 48, equation 2.5). The total waiting time passenger cost, which is the sum of the waiting time cost on each segment of the route, is calculated.

9. The in-vehicle passenger cost on each segment of the route is calculated. It is a function of the average running time on the segment and average load on the segment (during the hour of operation studied) weighted by the in-vehicle travel time cost. (see [4], pg. 50, equation 2.6). The total in-vehicle travel time passenger cost, which is the sum of the in-vehicle travel time passenger cost on each segment of the route, is also calculated.

10. The total passenger cost is the sum of the total waiting passenger cost and total in-vehicle passenger cost.

11. The scheduled operating cost is calculated. It is a function of the number of buses in operation, and the operating cost per hour. (see [4], pg. 50, equation 2.7).

The user can evaluate the cost of other schedules by changing the schedule parameters in step 1. Thus, the tool can be used to compare and evaluate alternative schedules by minimizing the total passenger cost.

**BASIS FOR THE RUNNING TIME MODEL**

The running time model proposed here is based on two key hypotheses and four additional assumptions:

1. The running time of a specific bus on a segment is independent of its headway with the immediately preceding bus

2. The running time of a specific bus on one segment is independent of its running time on the preceding segment

3. The passenger arrival rate is constant over a time period

4. All passengers can board the first bus

5. The schedule-based holding strategy is enforced

6. The scheduled segment running time does not affect the running time distribution

The first two hypotheses were tested using data from two routes of the Chicago Transit Authority (CTA) network, Routes 95E and 85 which will be used as case studies to evaluate the proposed scheduling process. Correlation analyses undertaken with the AVL data for these two routes during peak hours showed very small correlations between segment running times and headways, and between consecutive segment running times. The four additional assumptions were found to be reasonable.

**CHICAGO TRANSIT AUTHORITY ROUTE 95E APPLICATION**

This section investigates the potential for the improved generalised cost minimisation on bus Route 95E. Route 95E was chosen because of its apparent unreliability and because it is a “key route”, providing an extra incentive to improve reliability on this route.

---

4 The complete analyses can be found in pgs. 29-39 in [4].

5 Discussions of the assumptions are on pgs. 39-41 in [4].

6 Route 95E appears to have some segments with too much scheduled running time and others with not enough scheduled running time based on AVL data shown by the end-to-end running time analysis webpage on CTA intranet.
CTA Route 95E is a high-frequency bus route, which runs five miles East-West on 93rd St. and 95th St. from Buffalo St. to Lafayette St. Route 95 E connects with the south terminal of the Red Line, 95th/ Dan Ryan, two Metra stations and many bus routes serving either the Loop or the South part of the city. A schematic of Route 95E including the connections with the Red Line and the Metra Lines is shown in Figure 2. Operators begin and end their runs at 92nd St. and Buffalo St, the route’s eastern terminus. The CTA provides service throughout the day between Buffalo St. and Lafayette St. except between 23:30 and 4:30.

The characteristics of Route 95E are examined during the day and its reliability is assessed between 16:00 and 17:00. This hour is selected because it has the most boardings during the PM Peak. We are using one month of Automatic Passenger Count (APC) and Automatic Vehicle Location (AVL) weekday data, from September 25 to October 20, 2006 in this analysis. The main findings are the following: (1) There is low running time variability on the segments of the route—from the starting time point to arrival at the next time point. (2) The operators’ behaviour is inconsistent at time points; the variability of the holding time at time points is high, leading to unreliability. This is mainly due to too much running time scheduled on many segments. (3) There is an imbalance in trip time between the two directions; in the Westbound direction, the scheduled trip time is 31 minutes whereas the actual average trip time is 33.7 minutes, in the Eastbound direction, the scheduled trip time is 39 minutes whereas the actual average trip time is 31.6 minutes. Thus, 24% of the vehicles complete their Westbound trips within the allowed time whereas on average 90% of the vehicles complete their Eastbound trips within the allowed time. (4) Between 16:00 and 17:00, only 22% of the Westbound trips and 0% of the Eastbound trips depart on time. (5) The coefficient of variation of the headways varies between 0.58 and 0.86 showing high unreliability.

Before applying the model to Route 95E, the existing conditions were run on the model to confirm that the model produced results which are consistent with the data.

**Base Case: Strict Schedule-Based Holding Policy** - The model is then used first to provide a “base” case; the reliability of the current schedule is modelled assuming a strict

---

7 “Key” routes and “support” routes define the CTA bus system. Key routes provide the backbone of CTA service. They include the most productive bus routes, plus additional routes to provide basic geographic coverage. Support routes are the remaining routes. Key bus routes provide nearly half (47%) of all CTA rides. [CTA Service Standards]
8 The city center of Chicago.
9 For the full analysis, please refer to p. 55-62 of [4].
10 Since reliability problems propagate to following trips, the percentage of vehicles which depart on time from a terminal is different from the percentage of vehicles which completed the previous trip within the allowed time.
11 The analysis is on p.69 of [4].
schedule-based holding policy. The current schedule between 16:00 and 17:00 is assumed to operate throughout the operating day. The mean running times and headway variability are modelled between 16:00 and 17:00. This analysis allows the current schedule to be evaluated independently of current operator behaviour. We observe that between 16:00 and 17:00, after a day of operation, approximately 74% of the Westbound trips and 99% of the Eastbound trips are projected to be completed within the allowed time. Consequently, enforcing schedule-based holding allows a larger proportion of vehicles to start their next trip on time. Thus, 98% of the Westbound trips and 74% of the Eastbound trips depart on time. The analysis showed that schedule-based holding is critical to improving reliability for the waiting passengers along the route because it significantly lowers the variability of the headways compared to the current situation.

**Same Percentile Schedule** - The model was then used to evaluate the traditional ways to set vehicle scheduled running times; i.e., setting running times at a level where at least 50 or 65% of trips would have sufficient time to complete their trips on schedule. The analysis also intends to find the same percentile to implement on each segment in order to minimize the cost for the passengers while maintaining a reliable service. However, it was assumed that the recovery time will always allow at least 90% of the vehicles to run each trip within the scheduled half cycle time. We observed that the waiting time cost decreases slightly and the in-vehicle cost increases slightly with increasing percentile of the running time selected. However, the total passenger cost does not vary significantly across the full range of solutions. Implementing the 50th percentile on each route segment costs only 1.5% more in total passenger cost than the 10th percentile while yielding a more reliable service. Moreover, in reality, implementing the 10th percentile on each segment of the route would essentially mean eliminating all time points on the route. The same number of busses is needed to implement any schedule between the 10th and 75th percentile of the running time distribution. Consequently, the operating cost does not vary and stays constant for all running time distribution percentiles on this route. Therefore, if CTA chose to implement the same percentile on each segment of the route, one logical alternative would be to use the 50th percentile solution.

**Generalised Cost Minimisation Approach** - Finally, the model was used to obtain a schedule using the proposed generalised cost minimisation process detailed previously. The schedule obtained minimizes the total weighted customer minutes which is given by the sum of the passenger waiting minutes (weighted by 1.5) and the passenger in-vehicle minutes. The schedule obtained allows 97% of the trips in both directions to be completed within the allowed time and 97% of the trips in both directions would be able to depart on time. However, the scheduled segment running times obtained with the generalised cost minimisation approach are not reasonable for the operators; the approach is the same as suppressing each time point except 92 Com (92nd St. and Commercial St.). Indeed, we note that a very small proportion of vehicles are able to depart the segments on time, especially Eastbound. The schedule has been developed this way because on most segments of this route, more passengers are onboard than waiting for service, except at 92 Com. Consequently, the generalised cost minimisation approach sets low percentiles on the segments where there are more through passengers than waiting for service in order not to disadvantage a significant number of through passengers by holding at time points. Because there are few through passengers at 92 Com compared to the number of passengers waiting on the next segment, the only segment where a longer running time is proposed with the generalised cost minimisation approach is 92 Buf-92 Com. These results are consistent with the results of other researchers such as Liu [5].

**Cost Comparison** - Table 1 shows the costs for the passengers and for the CTA in dollars and in minutes for each schedule. The costs are calculated as explained in the previous section with the following cost parameters: Waiting time cost per passenger hour= $12/passenger-hour; In-vehicle travel time cost per passenger hour= $8/passenger-hour; Operating cost per hour of operation= $76/passenger-hour$^{12} (excluding pension costs). The fourth section of the table shows the excess waiting time, in-vehicle travel time and total

---

$^{12}$ Source: 2007 CTA Budget Recommendation p.139: Revenue Hours/Operating Cost= $76/hr of operation
The excess waiting time is that portion of the waiting time cost which is directly related to the variability of the headways on the route. If the route is perfectly reliable (i.e. the headways equal those scheduled), the excess waiting time will be equal to 0. The excess in-vehicle travel time is the time spent onboard by the through passengers when the vehicles are holding. The excess waiting and in-vehicle times presented in Table 1 are weighted by the number of passengers affected. The third and fifth sections of the table show the difference between each solution and the current situation for passenger minutes and excess passenger minutes, respectively, in percentage terms, a positive value indicating a lower cost for the alternative solution. First, we observe that the current schedule with schedule-based holding significantly improves reliability for the waiting passengers compared to the current situation. This further supports the view that passengers experience long waiting times on Route 95E primarily as a result of operator behaviour. We observe that any of the alternatives considered substantially improves reliability for waiting passengers on route 95E compared to the current situation. The passenger waiting times for each solution are much smaller than for the current schedule. In the 65th percentile approach and the current schedule with schedule-based holding enforcement, the through passengers are disadvantaged compared with the current situation because too much running time is scheduled. The schedules obtained with the generalised cost minimisation approach as well as the 50th percentile schedule have the lowest overall and weighted passenger costs. However, the excess time experienced by passengers with the two approaches is very different. Indeed, even though, the 50th percentile solution saves about half an hour (or 21%) of excess passenger waiting time compared to the generalised cost minimisation approach, it lengthens the through passengers in-vehicle time by over two hours or 50%. Overall, the generalised cost minimisation schedule saves a total of 112 (or 24%) of excess weighted passenger minutes compared with the 50th percentile schedule. In summary, the generalised cost minimisation schedule clearly shows more benefit for passengers compared to the 50th percentile scheduling approach and, in particular, for through passengers. Compared to the current situation, the generalised cost minimisation schedule saves 80% of excess total weighted time.

<table>
<thead>
<tr>
<th>Current situation</th>
<th>Current schedule with sch. Based holding</th>
<th>Traditional Approach (50%)</th>
<th>Traditional Approach (65%)</th>
<th>Generalised cost minimisation approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pax Waiting Time Cost</td>
<td>$572</td>
<td>$506</td>
<td>$506</td>
<td>$504</td>
</tr>
<tr>
<td>In-veh TT cost</td>
<td>$566</td>
<td>$613</td>
<td>$577</td>
<td>$590</td>
</tr>
<tr>
<td>Operating Cost</td>
<td>$532</td>
<td>$532</td>
<td>$532</td>
<td>$532</td>
</tr>
<tr>
<td>TOTAL COST</td>
<td>$1,670</td>
<td>$1,651</td>
<td>$1,615</td>
<td>$1,626</td>
</tr>
<tr>
<td>Waiting time in pax-min</td>
<td>2861</td>
<td>2530</td>
<td>2528</td>
<td>2519</td>
</tr>
<tr>
<td>In vehicle TT in pax-min</td>
<td>4244</td>
<td>4599</td>
<td>4327</td>
<td>4423</td>
</tr>
<tr>
<td>TOTAL PAX-MIN</td>
<td>7105</td>
<td>7130</td>
<td>6855</td>
<td>6942</td>
</tr>
<tr>
<td>TOTAL WEIGHTED PAX-MIN</td>
<td>8535</td>
<td>8394</td>
<td>8119</td>
<td>8202</td>
</tr>
<tr>
<td>% difference in waiting time</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
<td>11%</td>
</tr>
<tr>
<td>% difference in in-vehicle TT</td>
<td>-8%</td>
<td>-2%</td>
<td>-4%</td>
<td>2%</td>
</tr>
<tr>
<td>% difference in pax-min</td>
<td>0%</td>
<td>4%</td>
<td>2%</td>
<td>5%</td>
</tr>
<tr>
<td>% difference in weighted pax-min</td>
<td>2%</td>
<td>5%</td>
<td>4%</td>
<td>6%</td>
</tr>
<tr>
<td>Excess waiting time in pax-min</td>
<td>967</td>
<td>117</td>
<td>116</td>
<td>106</td>
</tr>
<tr>
<td>Excess in-vehicle time in pax-min</td>
<td>344</td>
<td>572</td>
<td>299</td>
<td>396</td>
</tr>
<tr>
<td>Excess total time</td>
<td>1311</td>
<td>688</td>
<td>415</td>
<td>502</td>
</tr>
<tr>
<td>Excess total weighted time</td>
<td>1795</td>
<td>747</td>
<td>473</td>
<td>555</td>
</tr>
<tr>
<td>% difference in excess waiting time</td>
<td>88%</td>
<td>88%</td>
<td>89%</td>
<td>85%</td>
</tr>
<tr>
<td>% difference in excess in-vehicle TT</td>
<td>-66%</td>
<td>13%</td>
<td>-15%</td>
<td>56%</td>
</tr>
<tr>
<td>% difference in excess pax-min</td>
<td>48%</td>
<td>68%</td>
<td>62%</td>
<td>78%</td>
</tr>
<tr>
<td>% difference in excess weighted pax-min</td>
<td>58%</td>
<td>74%</td>
<td>69%</td>
<td>80%</td>
</tr>
</tbody>
</table>

Table 1. Cost Comparison for Route 95E
CHICAGO TRANSIT AUTHORITY ROUTE 85 APPLICATION

The model was used for a second application of the scheduling method discussed previously to CTA Route 85. The analyses\(^{13}\) showed the same conclusion: the generalised cost minimisation approach schedule saved passengers travel time compared to the 50\(^{th}\) percentile schedule. It improves the reliability and the overall service quality for passengers on the route as long as schedule-based holding is enforced. The schedule obtained with the generalised cost minimisation approach saves 332 excess passenger minutes or 276 excess weighted passenger minutes compared to the 50\(^{th}\) percentile schedule. This represents a saving of 27% in excess total passenger minutes and 21% in excess weighted passenger minutes.

SENSITIVITY ANALYSIS

The sensitivity analysis sough to identify conditions under which the generalised cost minimisation scheduling process shows the most benefits over the traditional scheduling process of setting the 50\(^{th}\) percentile schedule on each segment of the route. The analysis showed that the generalised cost minimisation scheduling method is not sensitive to the total number of boardings on a route although it is sensitive to the ratio of waiting to through passengers. As the ratio of waiting passengers to through passengers increases, the generalised cost minimisation schedule shows greater benefits for waiting passengers compared to the 50\(^{th}\) percentile schedule. The analysis further demonstrated that the scheduling approach is sensitive to the location of the segment at which a high ratio of waiting to through passengers occurs. The schedule obtained with the generalised cost minimisation approach also showed benefits over the 50\(^{th}\) percentile schedule when the total number of waiting passengers on later segments is larger than the number of through passengers at the time point. Finally, the analysis showed that when the route is longer (unreliability propagates more easily on longer routes), the generalised minimisation schedule showed even more benefits over the 50\(^{th}\) percentile schedule. This is especially true when there were concentrated peaks of demand on the route.

CONCLUSIONS

This paper developed a scheduling process based on a model which explicitly projects and evaluates the tradeoffs between overall travel time and reliability. By simulating the running time distributions and headway variability of any proposed schedule, the model estimates the cost of the schedule for waiting passengers, onboard passengers and the transit agency. The scheduling process involves finding the time point schedule which minimizes the total cost with the help of the model. The scheduling process was applied to two Chicago Transit Authority bus routes; Route 95E and 85. For each route, the schedule which minimizes the total passenger cost was determined. The operating cost of the proposed schedule on each route is the same as for the current schedule because the same number of buses is used. The schedules obtained with the generalized cost minimization approach showed improved reliability and overall passenger service quality compared to the current schedule in both routes as well as compared to traditional approaches.

A sensitivity analysis showed that in most cases the generalized cost minimization schedule can significantly improve reliability and overall passenger service quality over traditional approaches.

REFERENCES


\(^{13}\) For the analyses, please refer to p. 102-126 of [4].