Dwell Time Model and Analysis for the MBTA Red Line

Andre Puong

03/30/2000

Abstract

High-frequency heavy rail operations are subject to performance variability, partly due to the impact of passsenger loads on the system. Specifically, vehicle dwell times at stations contribute to much of the running time and headway variations, constraining performance and service quality. Typically, dwell time is a function of passenger alighting and boarding volumes, and of the on-board crowding level. Based on observations made at stations on the MBTA Red Line, models showing linear effects in passenger boardings and alightings but nonlinear effects in the on-vehicle crowding level explains about 90% of the observed variation in dwell times. Furthermore, the nonlinear contribution of on-vehicle congestion is shown to be detrimental only to the boarding process. This lends to potential operations improvements if rationalized passenger movements can be achieved at high-volume stations.

1 Motivation

In any mode of public transportation, dwell time is a key parameter of system performance, service reliability and quality. Indeed, dwell time might represent a significant fraction of the total trip time along a serviced transit line, thus affecting travel time and system capacity.

In the case of heavy rail transit systems, where headways are short and running times between stations are roughly nonvarying across trains, dwell time is the main factor causing headway variability. Higher headway variability lowers service reliability in terms of on-time performance, and decreases service quality through longer waiting times and higher perceived on-board crowding levels.

Clearly, dwell-time is a function of some parameters that are not always controllable, such as door closing and opening mechanisms, vehicle loading conditions (which depends on both the train and platform configurations), and passenger arrival rates at stations. Nevertheless, *a priori* knowledge of dwell time at stations is most useful because:

- It gives insight into the travel time and headway variations over time of day. Combined with ridership information, it can yield effective timetables in order to maximize the throughput of the system.

- Efforts can be made at critical stations to enforce smaller dwell-time by modifying some of the abovementioned parameters. For instance, better-designed vehicles can lead to significant savings in vehicle loading and unloading times, two main components of dwell-time.

As an illustration, recent efforts have been made by the Metropolitan Tranportation Authority (MTA) of New York to reduce dwell times at Grand Central Station. Those efforts bore fruit as customer reaction was generally positive and rush dwell times improved by 10%, leading to a capacity increase of one train/hour (i.e., 2,000 customers during the peak hour).

In view of the important impacts of dwell-time on the overall performance of a rail transit system, we will attempt to derive dwell time models that show ease of use, consistency with respect to *a priori* considerations and predictability power. After a review of some current models and the theoretical aspects of dwell time modeling, a data collection procedure is described. Then a new model based on this data is presented and analyzed.

2 Theory and Current Dwell Time Models

Dwell time is defined as the time elapsed between the door opening and the door closing of a train sitting at a station. Clearly, in the absence of real-time operations (e.g., holding trains at stations), this time is devoted to the loading and unloading processes of the train, along with door opening and closing processes. Therefore, boardings and alightings at stations are likely the most significant factors causing dwell time variations.

We note that factors other than passenger loads and behavior also affect dwell times. Kraft [1] identified 7 major groups for these factors: human, modal, operating policies, operating practices, mobility, climate and other system elements. Yet, those factors are often constant in a given system or are beyond the knowledge/control of the operator. In consequence, given a particular rail system, those factors will be included in a dwell time model by

- grouping the system specific factors into a constant term.
- including the unknown factors in an error term.

Therefore, key factors for a straightforward model specification would be simply the number of passengers boarding and alighting, along with some measure of the crowding level in the train (e.g., arriving passenger loads). The first two variables are directly observable from the platform and can be collected with more or less resources depending on the station configuration and the ridership. The crowding level is harder to estimate and use because an accurate count of on-board passengers would require on-vehicle checks that are costly

while an approximate measure of the crowding level might not lead to a satisfactory model. Alternatively, if a new data collection is not affordable, accurate passenger boarding and alighting rates at stations (if known) could be used to infer passenger movements and loads.
We can first specify a simple model by assuming that the time T_{door} required to move passengers through a door increases independently and linearly with respect to the number of passenger boardings and alightings. This is equivalent to stating that the boarding and alighting phases are not interfering during T_{door}:

$$T_{door} = a + b \cdot B_{door} + c \cdot A_{door} \tag{1}$$

where

$$\begin{aligned} T_{door} &= \text{door open time,} \\ B_{door} &= \text{number of passengers boarding through door,} \\ A_{door} &= \text{number of passengers alighting through door,} \\ a, b, c &= \text{estimated parameters.} \end{aligned}$$

Adding a linear term containing the effect of congestion on-board would lead to:

$$T_{door} = a + b \cdot B_{door} + c \cdot A_{door} + d \cdot Congestion_{door} \tag{2}$$

where $Congestion_{door}$ reflects one or several of the following potential interferences:

1. The interference between alighting passengers and those staying on-board (through-standees)
2. The interference between boarding passengers and through-standees
3. The interference between the alighting and boarding passenger streams
4. The interferences among alighting passengers as their number increases
5. The interferences among alighting passengers as their number increases

Thus, the dwell time for a single car k can be taken as the longest T_{door} for all m doors of the car:

$$DT_k = max(T_{door\,1}, T_{door\,2}, \dots, T_{door\,m}) \tag{3}$$

Similarly, the total dwell time for an n-car train is the longest of the car dwell times:

$$DT = max(DT_1, DT_2, \dots, DT_n) \tag{4}$$

Given inputs about the passenger loads and movements at a station, we want to minimize the train dwell time at the station. Clearly, this minimum dwell time is reached when passenger loads and movements are evenly distributed over all doors and cars. In this case, the dwell time can be rewritten as follows:

$$\begin{aligned} DT &= \frac{a}{nm} + \frac{b}{nm}\,B + \frac{c}{nm}\,A + \frac{d}{nm}\,Congestion \\ &= \beta_0 + \beta_1\,B + \beta_2\,A + \beta_3\,Congestion \end{aligned} \tag{5}$$

where

B = total number of boardings,
A = total number of alightings, and
$Congestion$ reflects the overall on-board congestion.

Previous works (see [1] and [2]) have suggested that a simple dwell time model derived from (5) explains significant part of the dwell time variation when applied to non-congested situations:

$$DT = \beta_0 + \beta_1\,B + \beta_2\,A \tag{6}$$

In the case of the MBTA Green Line, T.M. Lin and N.H.M. Wilson have found in [2] that the model:

$$\begin{array}{llll} DT = 11.73 & +0.42 \cdot B & +0.49 \cdot A & (\bar{R}^2 = 0.68) \\ \quad\ (7.44) & \ (7.59) & \ (6.22) & \end{array} \tag{7}$$

explains about 70% (R^2) of the dwell time observations for the two-car trains on the Green Line, with all estimated coefficients significant (i.e., the t-statistics within parenthesis are all greater than 2.00). Several attempts were also made in [2] to include the car congestion effects on dwell time. One of the best models for the two-car trains was:

$$\begin{array}{lllll} DT = 13.93 & +0.27 \cdot B & +0.36 \cdot A & +8\,10^{-4} \cdot SUMASLS & (\bar{R}^2 = 0.70) \\ \quad\ (7.44) & \ (7.59) & \ (6.22) & \ (2.03) & \end{array} \tag{8}$$

where the last term is defined by

AS = number of arriving standees,
LS = number of departing standees, and
$SUMASLS$ = $A \cdot AS + B \cdot BS$

The congestion term attempts to capture the interfering effects of types (1) and (2) mentioned on page 3.

The new model presented in the next section is based on the results found in [2], but shows some significant improvements with respect to the effects of the number of door and the crowding on dwell times.

3 Study of the MBTA Red Line

3.1 The data collection procedure

A simple data collection procedure provided us for the necessary data to build a new dwell time model for the Red Line. Data was collected during Spring and Fall 1999 at Kendall and South Station on the Red Line, during the morning peak period. Except at Park Street station, only one platform is used for each direction of the line, thus only doors on one side were monitored to observe passenger flows.

Since we sought a dwell time model in the form of (5) (averaging passenger loads and movements over all doors and cars), only the total number of boardings and alightings, along with the total number of through standees were collected. Nevertheless, the latter was transformed so that the congestion level of the train was taken as the crowding level of the most congested cars. This is because those cars might lead to bottleneck during the boarding and alighting phases and drive the increase in dwell time.

Therefore, by only measuring dwell time, counting passengers movements through entrances and exits, and observing congestion levels (from the platform) of the most crowded cars the necessary data were obtained. Four people were needed for the data collection, each person being assigned several tasks (e.g., counting passenger movements through turnstiles). Clearly, depending on the labor resources available, counts at a car level and not at a entrance/exit level would be preferred for better accuracy.

3.2 The dwell time model

Given the prior analysis, we assumed that a simple and common dwell time model could be carried out for all types of trains (i.e., three and four-door car trains on the Red Line) and all stations where only one platform was used for loading and unloading (i.e., all but Park Street Station). Those assumptions were statistically tested[1] with the data collected at South Station and Kendall Station.
A total of 54 data points were used, combining three-door and four-door car trains. All dwell times were less than 1.5 min and passenger movements and loads showed enough variation to allow us to use an Ordinary Least Squares regression on our data. The statistical packages *SST* and *MINITAB* were used for the regression analysis.

The final dwell time function includes two major factors, the number of passengers boarding and alighting *per door*, plus one congestion term[2]:

[1] Details are given in Appendix B
[2] The hypothesis and analysis that lead to this model are discussed in Appendix A

$$DT = 12.22 \quad +2.27 \cdot B_d \quad +1.82 \cdot A_d \quad +6.2\,10^{-4} \cdot TS_d^3 B_d \quad (\bar{R}^2 = 0.89) \tag{9}$$
$$\qquad (12.82) \quad (7.11) \quad (9.07) \quad (4.70)$$

where

$$\begin{aligned} A_d &= \text{alighting passengers } \textit{per door}, \\ B_d &= \text{boarding passengers } \textit{per door}, \text{ and} \\ TS_d &= \text{through standees per door,} \\ & \quad \text{i.e., total through standees divided by the number of doors} \end{aligned}$$

The model fits 90% of the data: more precisely, 87% of the observations deviate from the value predicted by the model by less than 5 seconds. Besides, the congestion term can be easily interpreted in terms of interference between the boarding passengers and the through-standees[3]:

We can interpret the group of terms

$$2.27 \cdot B_d + 6.2\,10^{-4} \cdot TS_d^3 B_d = (2.27 + 6.2\,10^{-4} \cdot TS_d^3) B_d$$

as the marginal boarding time multiplied by the number of boardings (on average by door). If we plot this marginal rate against the number of through standees per door TS_d (see Fig. 1), we notice that the marginal boarding time

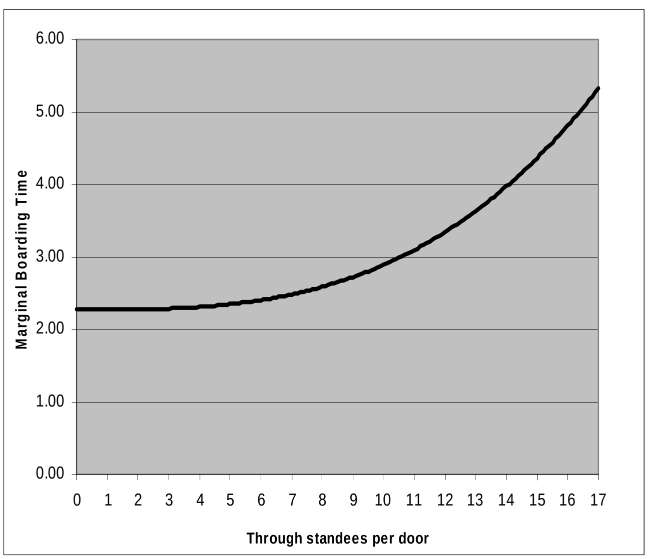

Figure 1: Marginal Boarding Time

[3]Refer to Appendix A for more details

through a door is an increasing function of the number of through-standees (i.e., the arriving standees who do not alight). This clearly meets our *a priori* condition about the effect of crowding on dwell-time variation: under 3 standees per door, the crowding effect is negligable but beyond this threshold, an ”incremental” standee per door worsens significantly the boarding process. Neither on-board standees nor the passengers on the platform seem to affect the alighting process or at least to only a minor extent. This might be explained by the discipline of the passengers who alight and board without interfering with each other. Therefore, the bottleneck that drives the dwell time variation is the on-board congestion that precludes fast boarding. Note that this adverse effect is all the more important as most of the boardings may occur at the already congested cars. Our model actually overestimates the marginal boarding time since average boardings are used to explain the dwell time variation.

4 Operational Implications

The analysis performed on the data collected at Kendall and South Stations show that dwell time, which consist of the time needed for all passengers to alight and board, is highly sensitive to the presence of standees in the car. This suggests that, in the absence of measures to improve the alighting and boarding processes on the platform, on-board congestion and boardings can interact negatively so that dwell time increases rapidly between successive stations. This leads in turn to train bunching and deterioration of service quality (longer waiting times at stations and crowding conditions in the train).
Past efforts by transit authorities at London, Hong Kong and New York City have shown that dwell time reduction thanks to safer and more orderly boarding results in improved throughput.

5 Conclusion

This research has emphasized the variability of dwell time with respect to passenger movements and loads. A simple model, which is accurate and consistent with *a priori* considerations, has been estimated and has underlined the effect of on-board congestion on boarding times for the non-Park Street Stations of the MBTA Red Line. Moreover, this effect was shown to be non-linear: it was better approximated by a polynomial function of the number of standees.
This sensitivity of dwell time to ridership variations exposes train operations to uneven headways and running times, leading to a deterioration in service quality and capacity. It is believed that the model presented lays the basis for a better identification of the critical stations for maintaining high-frequency service during peak periods.

A Model Specifications

We tried to represent as accurately as permitted the mechanisms that drive the four components of the dwell-time, i.e. door opening time, aligthing time, boarding time and door closing time. In comparison with the models developed by T.M. Lin in [2] (see equations (7) and (8)), the model presented in equation (9) captures some new effects described below:

1. The difficulty encountered by T.M. Lin to model dwell times for multiple-car trains is circumvented by dividing our data by the total number of doors. This means that we consider the case where movements and loads are evenly distributed. Although this is unlikely the case for most stations, this model specification gives us a lower bound on the dwell time.

2. In congested situations, we can realistically represent the additional dwell time by considering the variation of the marginal alighting and boarding times with respect to the crowding level.
   In the case of the boarding process, we suggest that the increment of marginal boarding time is a function of the number of standees remaining in the vehicle after the alighting process (through standees). This formulation accounts for instance for the increasing difficulty of boarding with the congestion level of the car.
   Therefore, an additional component $g(TS_d) \times B_d$ can be added to the dwell-time equation, where g is a piecewise increasing function representing the additional average marginal boarding time. Under a threshold value of TS_d^{crit} , there is no interaction between boarding passengers and standees, so that $g(TS_d) = 0$ for $TS_d \leq TS_d^{crit}$. For $TS_d \geq TS_d^{crit}$, f is an increasing function of TS_d.
   Similarly, we can consider an additional term $f(AS_d) \times A_d$ accounting for the interaction of standees -when the train arrives at the station- with the aligthing passengers, where:

   AS_d is the number of arriving standees per door,
   and f is a function with a shape similar to g (see figure 2 below).

We approximate our functions f and g by two polynomial functions of order 3. Thus,

$$\begin{aligned} DT =& \beta_0 + \beta_1 A_d + \beta_2 B_d + \beta_3 TS_d \times B_d + ... \\ &+ \beta_6 TS_d^3 \times B_d + \beta_7 AS_d \times A_d + ... + \beta_{10} AS_d^3 \times A_d \end{aligned} \tag{10}$$

A stepwise regression was performed and lead us to the simpler form presented on page 6:

$$DT = 12.22 + 2.27 \cdot B_d + 1.82 \cdot A_d + 6.2\,10^{-4} \cdot TS_d^3 B_d \tag{11}$$

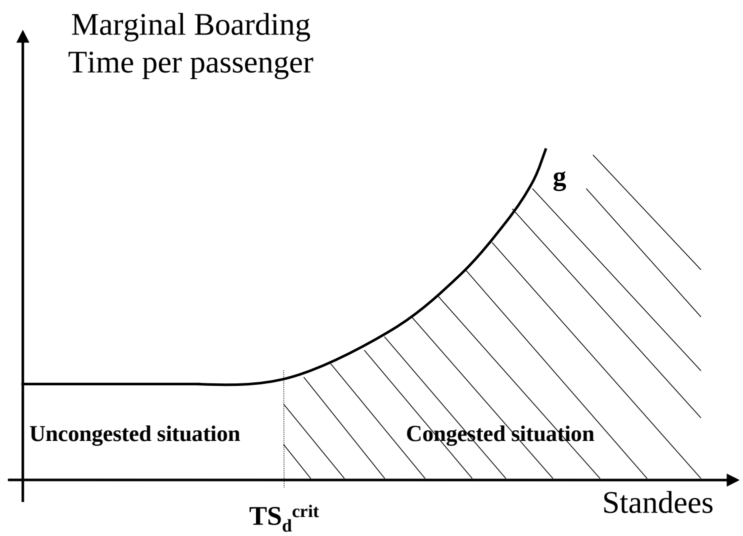

Figure 2: Marginal Boarding Time

B Statistical Tests of The Model

1. **Test of a Single Model for Kendall and South Stations**
   A Chow-test was performed to ensure that the model defined by equation (9) can apply to the two different datasets collected respcetively at Kendall and South Station. The result of the test validates the hypothesis that we can use a single model for different stations of the Red Line (except for Park Street Station).

2. **Test of Serial Correlation**
   The Durbin-Watson statistic given in Table 2 (2.37539) does not indicate serial correlation since $4 - d_u \leq 2.37539 \leq 4$, which indicates uncertainty at a 95% level of confidence.

3. **Test of Heteroscedaticity**
   No presence of heteroscedasticity was shown by a White test.

C Data and Statistics

Table 1: Data Statistics

	DT	A_d	B_d	TS_d^3B
Mean	28.91228	3.42106	3.59966	2.27069e+003
Standard deviation	13.15469	2.72049	2.99871	4.88231e+003
Minimum	13.00000	0.50000	0.66670	0.00000e+000
Skewness	1.48377	0.98774	2.4134	2.91847
Maximum	77.00000	10.08330	17.44440	2.28061e+004

Table 2: Regression Statistics of equation (9)

Dependent Variable: DT

Independent Variable	Estimated Coefficient	Standard Error	t-Statistic
1	12.22254	0.95346	12.81913
A_d	1.82224	0.25623	7.11170
B_d	2.26925	0.25010	9.07354
$TS_d^3B_d$	6.20604e-004	1.31822e-004	4.70791

Number of Observations	54.
R^2	0.89981
$\bar{R}^2$	0.89380
Sum of Squared Residuals	8.16318e+002
Standard Error of the Regression	4.04059
Durbin-Watson Statistic	2.37539
Mean of Dependent Variable	27.75926

References

[1] W.H. Kraft. *An Analysis of the Passenger Vehicle Interface of Street Transit Systems with Applications to Design Optimization.* Ph. D. dissertation. New Jersey Institute of Technology, 1975.

[2] T.M. Lin, N.H.M. Wilson. *Dwell Time Relationships for Light Rail Systems.* Transportation Research Record 1361, 1991.