Australasian Transport Research Forum 2017 Proceedings
27 – 29 November 2017, Auckland, New Zealand
Publication website: http://www.atrf.info

Impact of the Spatial and Temporal Variation in Passenger Service Rate on Train Dwell Time

Sunhyung Yoo[1], Dr Jinwoo (Brian) Lee[2], and Dr Hyun Kim[3]

[1,2]School of Civil Engineering and Built Environment, Science and Engineering Faculty

Queensland University of Technology (QUT), Brisbane, Australia

[3]Department of Railway Research, Korea Transportation Institute, Republic of Korea

Email for correspondence: brian.lee@qut.edu.au

Abstract

This study aims to improve the understanding of the underlying mechanism of passenger boarding and alighting processes, as well as its potential influence on train dwell time. The study examines if the interference between different passenger movements (boarding, alighting, and standing) affects the length of passenger service time. This study also postulates that the level of interference can be related to the position of the train door in terms of its proximity to platform entries. We collected passenger service time data from one of busiest metro stations in Seoul, Korea, and analysed the empirical data to identify the spatial and temporal variation in the passenger service rate (number of passengers serviced at each door per unit time). This study introduces the Dynamic Time Warping technique for similarity measures and clustering of the time series data. The analysis revealed four distinct shapes of the temporal service rate curve. Each curve represents a unique temporal service pattern that is attributed to varying levels of interference and passenger demand. A new passenger service time model is presented to include the cluster variable in a categorical form. The significance of the cluster variable is demonstrated by comparing the prediction capability of the model with and without the variable.

Keywords: train dwell time, passenger service time, dynamic time warping

1. Introduction

Train dwell time is composed of two components: function time and passenger service time. Function time is a buffer to allow for door opening, safety check before door closing, and door closing. Passenger service time refers to the period from the door opening to the completion of passenger service. Passenger service time varies from one train door to another, and the dwell time of a train is determined by the longest passenger service time. Passenger service time is exposed to a broad range of variations. The number of boarding passengers and the number of and alighting passengers are the dominant factors in determining the length of passenger service time (Szplett, 1984; Lam et al., 1998; Wiggenraad, 2001). Weston (1989), Lin and Wilson (1992), Puong (2000), and Douglas (2012) included an interference factor in various terms to account for the disruption to passenger boarding and alighting. Other studies take into account the variation in train vehicle and station types by including door width and train length factors (Harris, 1994; Li et al., 2015).

The majority of train dwell time studies in the literature assume the static passenger service rate—in other words, the numbers of boarding and alighting passengers per unit time are constant over the whole passenger service time. Harris (2006) argued against this common assumption. His empirical study of the London metro data shows substantial variations in the passenger service rate over time and among train door locations. Several factors may be responsible for these variations. In particular, passengers are unequally loaded onto train cars. Waiting passengers tend to position themselves near the platform entrances (e.g., stairs or escalators) rather than being distributed equally (Szplett and Wirasinghe, 1984; Krstanoski, 2014). Unevenly distributed passengers contribute to varying levels of interference both spatially and temporally.

This study is premised on the hypothesis that the passenger service rate varies over time and among train doors, and that the position of train doors in terms of its proximity to platform entrances is a major influence on the variation. Empirical data were collected through a field study of the busiest metro stations in Seoul, Korea. The data are essentially a time series representing the passenger service rate (i.e., the number of boarding and alighting passengers in every 5-second interval) at each door location. The empirical data are analysed to demonstrate the heterogeneity in the service rate and to identify potential sources of the interference. This study introduces the Dynamic Time Warping (DTW) technique for the similarity assessment and clustering of train doors by their temporal service pattern. A passenger service time model is presented to include the cluster information as a categorical variable. The significance of the temporal service pattern and its effect on the passenger service time are tested by comparing the prediction capability of the model with and without the cluster variable.

The next section reviews the literature on train dwell time and passenger service time models. The following section describes the study site and the empirical data collected for the study. The data are analysed to reveal spatial and temporal variation for similarity measures and clustering using DTW. The clustering results are presented with discussions. The last section presents a passenger service time model to test the significance of the cluster variable. The modelling results are discussed, as well the main findings of the study and their implications on railway operation.

2. Train dwell time models

The literature on train dwell time or service time has been scarce, with few advances since the 1980s. Recently, passenger behaviour is receiving attention to address the overcrowding issue of urban metros. Some studies examined passenger behaviours to better understand temporal boarding and alighting processes (Krstanoski, 2014), as well as passengers' choice of waiting position on the platform (Kim, 2014) or a specific train car (Liu et al., 2016). Ahn et al. (2016) proposed that passenger concentration on the platform could be reduced by providing passengers real-time train arrival and occupancy information. Coxon et al. (2013) also suggested that train performance could be improved by refining the interior design and seat configuration to allow smooth and stable passenger movements in train cars.

Train dwell time may be defined as the time duration from the moment of train arrival at the station to the moment of departure, or, in other words, the time duration while a train stops in the station. The literature defines train dwell time as the sum of two components: function time and passenger service time (Lam et al., 1998). Passenger service time is the period of passenger boarding and alighting. The dwell time models in the literature attempt to estimate the passenger service time component.

Earlier dwell time models include only the most dominant determinants of passenger service time, namely the number of boarding passengers and number of alighting passengers. Wirasinghe and Szplett (1984) proposed a linear regression model using the empirical data collected from two Light Rail Transit (LRT) stations in Calgary, Canada. Their model was designed to estimate the passenger service time for each train door. The maximum serivce time out of all the doors was then added to the function time to compute a train's dwell time. A similar model was proposed by Lam et al. (1998) for passenger service times at Mass Transit Railway stations in Hong Kong. The model assumed spatially uniform distribution of boarding and alighting passengers across the train doors. More recently, Jong and Chang (2011) proposed two regression models for the commuting and express train lines at the Taipei Main station in Taiwan to reflect the impact of different train types.

The majority of existing dwell time or passenger service time models assume that passenger boarding and alighting rates are constant over the whole passenger service time. Harris (2006) challenged this common assumption and demonstrated that the rates of boarding and alighting vary widely through a cohort of passengers. His empirical study showed that, in general, the alighting rate is fastest immediately after doors open and that the fastest boarding rate is found with the middle third of passengers; early boarders can be impeded by those alighting, and late boarders may be impeded by the congestion in the train and station terminal.

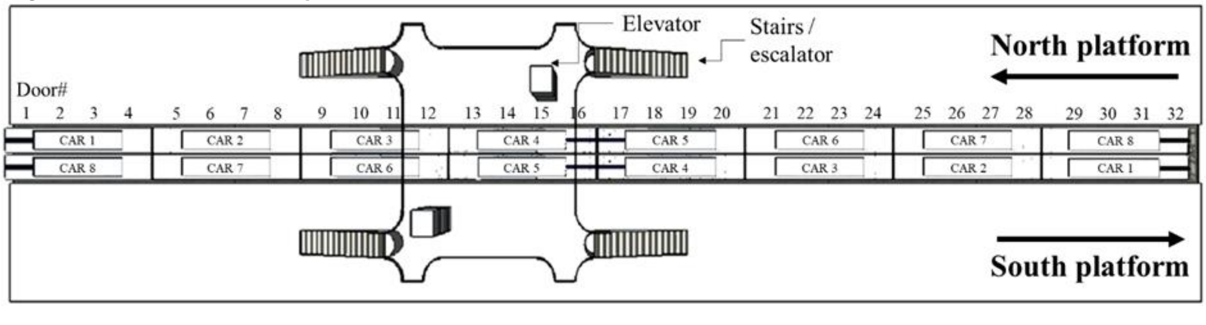
The spatial distribution of waiting passengers along the station platform is a major factor contributing to passenger service interference and service time variation. Szplett and Wirasinghe (1984) and Wiggenraad (2001) showed that train passengers unevenly distribute themselves along the platform and found that the spatial distribution of waiting passengers is highly related to the position of the platform access points. More recently, a number of models were proposed to explain the spatial distribution of passengers across the station platform. Kim (2014) used a logit model to determine the factors affecting the selection of a waiting position on the platform. Krstanoski (2014) suggested multinomial distribution to describe the passenger boarding and alighting process over time. This study also found a clear dependence of the spatial distribution of passengers on the position of the platform entrances. Liu et al. (2016) proposed a utility-based model to describe the choice of a specific car on a train. Different modelling approaches are discussed in the literature to explain the positioning and distribution pattern of passengers. However, the consequences and impact of the spatial and temporal distribution of passengers on passenger service time has been unexplored and represents a gap in the current literature.

3. Study site and data

To collect passenger service time and associated data, a field study was conducted using a video camera to capture passenger boarding and alighting activities. The Gasan Digital Complex (GDC) station in Seoul, Korea, was selected for the study. The GDC station is one of busiest metro stations in Seoul; more than 50,000 passengers accessed the station per day in 2015. Severe dwell time delays are seen in this station that substantially exceeds the scheduled dwell time, especially during the afternoon peak hours. This study uses train car weight data as a proxy measure of the crowd level in a train car. Seoul Metro currently operates a weighing scale system at every metro station, which enables the collection of the weight data of each train car on the real-time basis.

The GDC Station has a side platform structure (Figure 1) providing two separated platforms: one towards Bu-pyeong station (north platform) and the other towards Jang-am station (south platform). This study uses only the north platform data, which is the main travel direction during the afternoon peak-hours. The metro line serving this station operates eight car trains, and there are four entry/exit doors on each side of every train car. Figure 1 illustrates the designated positions of the 32 train doors and the location of passenger access points—stairs, escalators, and elevators.

Figure 1: GDC station layout



The data collection took place on April 29, 2015, from 6 to 8 p.m. on north platforms. A total of 16 on the north platform were installed so that each camera could capture passenger boarding and alighting movements and associated activities at each door location. The following data were derived from the video recordings and also from the weighing scale system:

- Weight of train cars – The weight measurements of each train car after completion of passenger service between door closing and train departure were recorded.
- Function time – Function time is defined as the time period from train arrival to door opening and the time period from door closing to train departure. The function time is uniform for all the train doors.
- Passenger service time – Passenger service time is defined as the time period from the instance of the door opening to the instance of the last passenger's boarding or alighting. The passenger service time was measured at each train door location.
- Number of boarding and alighting passengers – The number of boarding and alighting passengers was counted separately in 5-second intervals at each train door location.
- Arrival and departure time of trains – The arrival time and departure time of trains were recorded for every train arriving at GDC station. Train arrival time was recorded when the train stops completely. Train departure time was recorded when the train begins to move away from the station.

The passenger service time data were collected from a total of 27 northbound trains during the study period. Out of those 27 trains, the scale system failed weighing one train. Additionally, one video camera installed on the north platform to capture door #13 and #14 failed to record. Therefore, the service time data were collected from 26 northbound trains at 30 door locations.

Table 1 presents descriptive statistics of the variables of interest. The average gross weight of the northbound train cars is 9.95 tons. The tare weight of a train car is approximately 0.5 tons. These measurements of average gross weight indicate that the average passenger occupancy for the northbound train cars is approximately 155 passengers. The maximum occupancy in a car is about 240 passengers, assuming the average weight of male and female adults is 71.5 kg and 57.0 kg, respectively (KOSIS, 2013).

On average, there are 12.11 boarding passengers and 3.97 alighting passengers per door per dwell. The number of boarding passengers is relatively consistent, with less variation compared to the alighting passengers. The mean passenger service time is 15.27 seconds with a standard deviation of 7.54 seconds. The average function time is 13.74 seconds. The function time includes the pre-service function time (time period from train arrival to door opening) and the post-service function time (time period from door closing to train departure). The post-service function time is typically longer than the pre-service function time because it involves the safety check by the train operators to ensure that all the train doors are safely closed before departure. The post-service function time was also found to be more variable compared to the pre-service function time.

Table 1: Descriptive Statistics

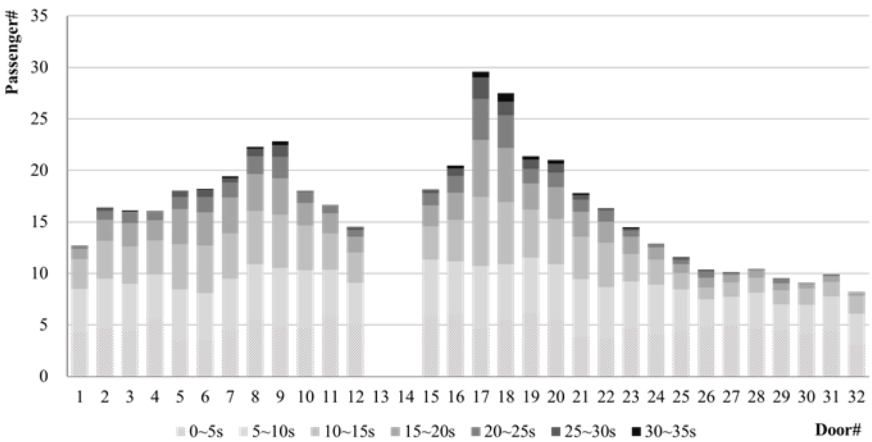
Type of Data		N	%	Mean	Std. Dev.	Coeff. of Variation
Train car weight (WT)	**0-5 tons**	7	0.97%	9.95	3.32	0.33
	5-10 tons	405	56.41%			
	10-15 tons	237	33.01%			
	over 15 tons	69	9.61%			
Number of boarding passengers per door	**0-5**	135	18.80%	12.11	6.85	0.57
	6-10	192	26.74%			
	11-15	185	25.77%			
	over 16	206	28.69%			

Type of Data		N	%	Mean	Std. Dev.	Coeff. of Variation
Number of alighting passengers per door	**0-5**	547	76.18%	3.97	3.33	0.84
	6-10	135	18.80%			
	11-15	30	4.18%			
	over 16	6	0.84%			
Passenger Service time (sec)	**0-5**	65	9.05%	15.27	7.54	0.49
	6-10	155	21.59%			
	11-15	168	23.40%			
	over 16	330	45.96%			
Function time (sec)	**Pre-passenger service**	27	N/A	5	0.73	0.15
	Post-passenger service	27	N/A	8.74	2.8	0.32

4. Passenger service time analysis

The service time data is analysed in this section to demonstrate spatial and temporal variation in association with the level of passenger flow. Figure 2 illustrates the average passenger service rate, which was counted at each train door location during the study period. Note that the service rate is expressed in terms of the number of boarding and alighting passengers serviced at each door location in 5-second intervals. The overall height of the bars indicates the total number of passengers serviced per passenger service time.

Figure 2: Passenger service rate in 5-second intervals by train door



The spatial variation of the passenger load is substantially large among the door locations. An average of 29.6 passengers used door #17 per passenger service time, whereas only 8.2 passengers used door #32. Figure 2 shows that passengers tend to cluster around certain designated locations rather than evenly spread along the platform. Two cluster locations were found at each platform—around door #17 and #9. The cluster locations are close to the platform entry points (Figure 1).

Figure 2 shows the passenger service rate in 5-second intervals. The service rate in the first 5-second interval varies substantially, ranging from the lowest service rate of 3.1 passengers (door #32) to the highest service rate of 6.1 passengers (door #16).

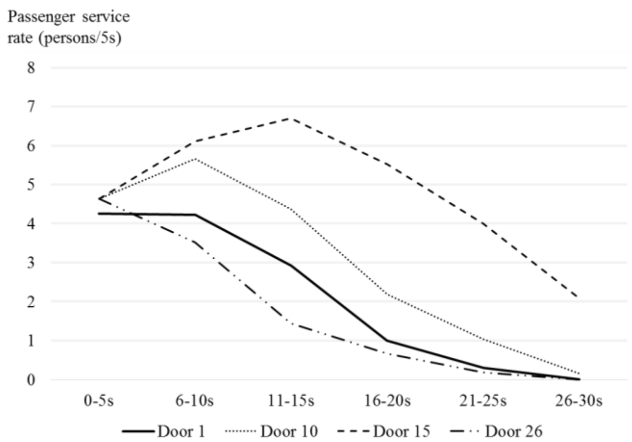
Table 2: Passenger service rate variation

Time Interval (sec)	Number of serviced passengers		
	Mean	Std. Dev.	Coeff. Of Variation
0~5s	4.67	0.80	0.17
5~10s	4.55	1.02	0.22
10~15s	3.41	1.50	0.44
15~20s	2.13	1.34	0.63
20~25s	1.09	0.88	0.81
25~30s	0.50	0.49	0.98
30~35s	0.23	0.22	0.93
35~40s	0.09	0.02	0.23

As shown in Table 2, passenger service rates are relatively high and less varied in the first and second intervals. Afterwards, the variation becomes larger—most of the waiting passengers have boarded the train already at some door locations, but the passenger queue was also still present at other door locations. The service rate should gradually decline as waiting passengers board the train, but an interesting trend was found at some door locations, in which the passenger service rate further increased in subsequent time intervals. For instance, the service rate at door #17 continues to increase from 4.6 passengers in the first interval to 6.1 passengers in the second interval, and then to 6.7 passengers in the third interval.

Figure 3 shows some examples of the temporal passenger service rate pattern (door #1, #10, #15, and #26). The profiles show a general agreement in the trends, but the peak service rate point varies by the door location. The highest service rates are seen in the first 5 seconds at door #1 and #26, whereas the service rate is at its peak in the second interval at door #10 and in the third interval at door #15.

Figure 3: Temporal passenger service rate patterns (door #1, 10, 15 & 26)



The temporal service rate patterns illustrated in Figure 3 are consistent with the discussions in the literature in that passenger movements at the beginning of service time can be substantially disrupted by interaction among passengers (Weston, 1989; Lin and Wilson, 1992; Parkinson and Fisher, 1996; Puong, 2000; Douglas, 2012). Initial passenger boarding and alighting can be drastically slowed or impeded by standing passengers in the train and waiting passengers on the platform (Krstanoski, 2014). This interference can be more severe when conjoined with loaded train cars and a concentration of waiting passengers. The service rate rebounds after the interference becomes less obtrusive. The level of interference can be associated with the occupancy level of the train vehicle and the position of stairs and elevators, as waiting passengers tend to cluster near the platform entry points.

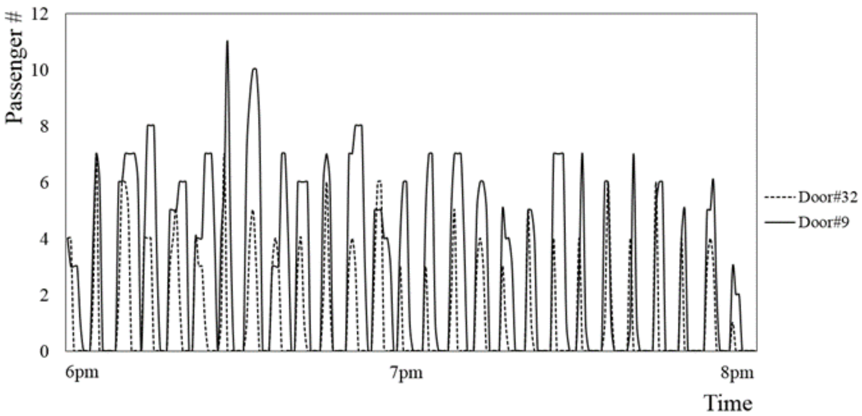
5. Time series clustering

5.1. Dynamic time warping algorithm

The passenger service rate data are further examined for similarity assessment and clustering in this section. In this paper, we introduce DTW to compute the degree of similarity in the service rate pattern between door locations. For time series clustering and similarity assessments, the Euclidean norm was widely used in spite of its sensitivity to distortion in the time axis (Keogh and Kasetty, 2002). DTW allows a non-linear warping of the time axis to find an optimal alignment between X and Y with the least cumulative cost, called the warp path (Keogh and Pazzani, 2001). DTW was originally proposed for automatic speech recognition (Rabiner and Juang, 1993), but it has been applied to a wide range of applications for data mining (Berndt and Clifford, 1994), information retrieval (Chu et al., 2002; Keogh and Ratanamahatana, 2005), bioinformatics (Gavrila and Davis, 1995; Munich and Perona, 1999; Kovacs-Vajna, 2000; Aach and Church, 2001), chemical engineering (Gollmer and Posten, 1995), robotics (Schmill et al., 1999), and computer graphics (Caiani et al., 1998).

The empirical data are presented as time series with different lengths. Passenger service time data were aggregated into a time sequence for a two-hour interval for the analysis. Figure 4 illustrates two samples of aggregated data showing different temporal patterns. Door location #9 is one of busiest waiting areas on the north platform, in which the number of passengers is substantially higher than those at door location #32.

Figure 4: Sample aggregated service time sequences (door #9 & #32)



Suppose that two service time datasets (each collected at a different door location) are given as $X = \{x_1, x_2, \dots, x_n\}$ and $Y = \{y_1, y_2, \dots, y_m\}$. Euclidean distance is simply the sum of $|x_1 - y_1|, |x_2 - y_2|, \dots,$ and $|x_n - y_m|$ by calculating the difference between two points occurred at the same time (i.e., local distance). The DTW algorithm starts by calculating the local distance between two parallel points (with the same time tag) x_i and y_j. The local distance is calculated as follows:

$$dist(x_i, y_j) = |x_i - y_j| \quad \textbf{(1)}$$

The local distance value between two points is small if x and y are closely positioned to each other, and the distance is large if x and y are far apart. A *m*-by-*n* cost matrix can be obtained by computing the local distance between all elements of X and Y; the (i^{th}, j^{th}) element of the matrix contains the distance between the two points x_i and y_i.

Once the calculations of local distance for sequences X and Y are obtained, the next process is to find the optimal alignment between X and Y to have the minimum overall distance. This process requires the calculation of the cumulative cost matrix $D(i, j)$. The naive approach for alignment is to try every possible warping path between X and Y, leading to extensive computational complexity. Instead, a dynamic programming algorithm is widely used to find the shortest warping path, which allows the algorithm to move to the next cell vertically, horizontally, or diagonally to minimise the accumulated cost (Rabiner and Juang, 1993). The accumulated cost can be calculated using the following (2):

$$D(i,j) = dist(x_i, y_j) + \min\big(D(i-1,j), D(i-1,j-1), D(i,j-1)\big) \quad \textbf{(2)}$$

Next, the optimal path allocates the alignment for DTW distances between X and Y. The optimal path may be defined as sequence $p = \{p_1, \dots, p_l, \dots, p_L\}$ with $p_l = (n_l, m_l) \in [1{:}N] \times [1{:}M]$ satisfying the following three conditions:

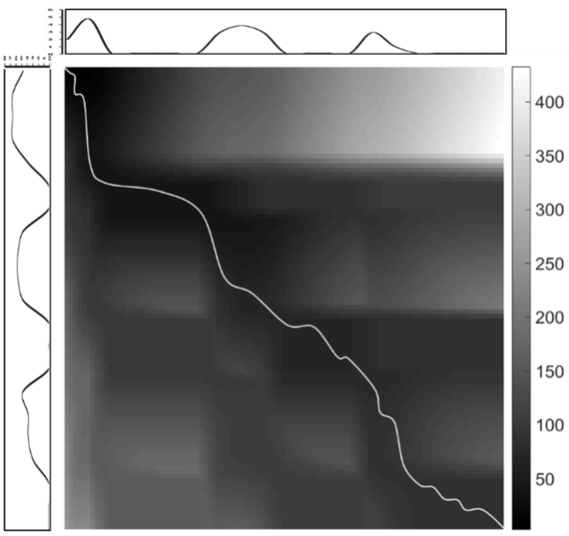
(i) *Boundary condition*: $p_1 = (1,1)$ and $p_L = (N, M)$. This constraint simply defines that the warping path has to begin and to end in diagonally opposite corners of the matrix. For the alignment, this condition implies that the first and last points of the two time series have to be aligned to each other.

(ii) *Monotonicity condition*: $n_1 \leq n_2 \leq \cdots \leq n_L$ and $m_1 \leq m_2 \leq \cdots \leq m_L$ which means that the path should move down or to the right from the p_1 in the accumulated cost matrix.

(iii) *Step size and continuity condition*: $p_{l+1} - p_l \in \{(1,0), (0,1), (1,1)\}$ for $l \in [1{:}L-1]$. This constraint defines that only adjacent cells (vertical, horizontal, or diagonal) can be selected for the next step and also that the skipping of cells is not allowed.

Finally, the DTW similarity measure is an accumulated cost of the last value of the optimal path. Comparing the two service time series of door locations #9 and #32 on the north platform, the DTW similarity measure (p_L) was calculated as 85 (Figure 5). The smaller the DTW similarity measure is between two sequences, the more similar these sequences are. By using this DTW algorithm, this study endeavoured to calculate all DTW similarity measures between the time series data of all train doors for clustering.

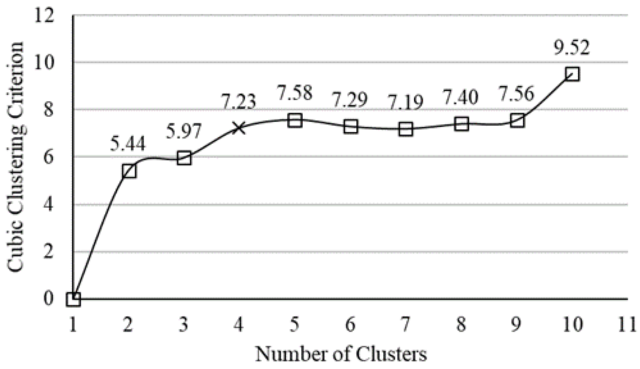
Figure 5: Cost matrix and optimal path (door #9 & #32)



5.2. Train doors clustering

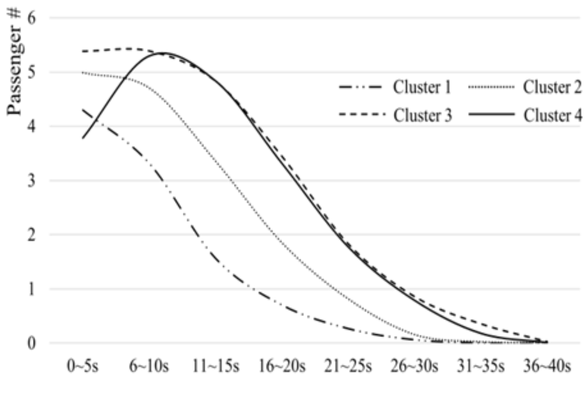
The cubic clustering criterion (CCC) was used to determine the number of clusters. The CCC is a function of the ratio between the R^2 with a specific number of clusters and the expected R^2 with a uniformly distributed set of events (Sarle, 1983). Through an extensive Monte Carlo experiment, Sarle (1983) proposed that a CCC value greater than 2 or 3 would be sufficient evidence for the validity of the chosen number of clusters. In this study, the optimal number of clusters was chosen by plotting the CCC value against the number of clusters and finding a local maximum after the CCC rises above 2 and the rate of change becomes steady. Four clusters were created as a result of the clustering.

Figure 6: Cubic clustering criterion plot



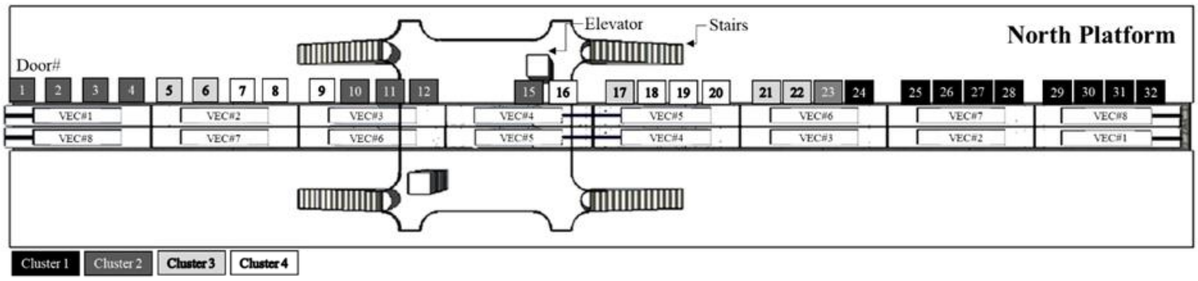
The temporal passenger service pattern is illustrated in Figure 7 for each cluster group. The curves clearly demonstrate uneven temporal patterns by the location of train doors. The interference effect is most significant in the service pattern of cluster 4 (solid curve). The same effect is present in the curves of cluster 3, but the degree of interference is relatively small and the service rate only slightly increases compared to the cluster 3. The plots indicate that the level of passenger demand determined the temporal service pattern for cluster 1 and cluster 2, in which the interference effect is not observed.

Figure 7: Cluster dendrogram



The clustering result is visualised on the station layout in Figure 8. The figure confirms that the location of doors is significantly related to the shape of the curves and the level of interference. Most train doors in cluster 4 are located closest to the platform entry points. Heavy boarding and alighting passenger volumes may have caused significant interference and disruption to passenger movements at those door locations. The doors in cluster 3 are mostly located next to the platform entry points. The walking distance to the entry points gradually increase from the doors in cluster 3 to cluster 2, and from cluster 2 to cluster 1. The train doors categorised in cluster 1 are located on the corner of the platforms.

Figure 8: Clustering result displayed on the station layout



6. Passenger service time model

To examine the validity of the temporal service rate pattern and its impact on passenger service time, a linear regression analysis was conducted. The temporal service pattern is included in the models as a categorical variable representing the cluster number. The proposed models consist of the passenger service time as a dependent variable and four explanatory variables, which includes the number of alighting passengers, the boarding of boarding passengers, the weight of train car, and the train door cluster variable. Three of the independent variables are continuous and quantitative variables, so they can be analysed without any modifications. The train door cluster variable is a categorical variable so it was transformed into dummy variables.

Table 3 presents the modelling results with and without the cluster variable. The table includes unstandardised regression weights, their standard errors, and standardised beta weights (β). The model without the cluster information was statistically significant with adjusted R^2 at .928. Including the cluster variable improved the prediction capacity of the model. The adjusted R^2 value improved when the clustering variable was included in the models, and all the explanatory variables were statistically significant with 99% confidence. The results suggest that the temporal service pattern has a significant impact on passenger service time, and the linear models explain the relationship reasonably well.

Table 3: Regression Modelling Results

	Without Clustering			With Clustering		
Variables	***B***	***SE B***	***Beta (β)***	***B***	***SE B***	***Beta (β)***
	Adjusted R^2=.928, $F(3,714)$=3077.14, p<.0001			Adjusted R^2=.937, $F(6,711)$=1782.31, p<.0001		
Constant	-.302	.273		1.598***	.332	
No. of Alighting Passengers	.885***	.029	.346	.769***	.029	.301
No. of Boarding Passengers	.900***	.015	.725	.857***	.015	.691
Train Vehicle Weight (Ton)	.153***	.029	.060	.212***	.028	.083
Cluster_1				-2.344***	.261	-.129
Cluster_2				-2.207***	.252	-.117
Culster_3				-.717***	.259	-.032

Note: *B* = unstandardised regression weights; SE B= standard error for the unstandardised regression weights;$^{***}p$<.001; $^{**}p$<.01; $^{*}p$<.05

The number of boarding passengers and the number of alighting passengers are the dominant factors determining the passenger service time. The passenger service time increases by 0.769 with an additional alighting passenger and by 0.857 seconds with an additional boarding passenger on the north platform. The train car weight is the next significant variable, in which the value of β is .083. The unstandardised coefficients (B) of dummy variables imply the relative level of interference delaying passenger movement and service time. The interference is most crowded at the door locations in cluster 4, followed by cluster 3, cluster 2, and cluster 1. These negative signs of dummy cluster parameters indicate that the door locations in cluster 4 would take longer passenger service time than the door locations in the other clusters.

7. Conclusions

To better understand the underlying mechanism of passenger boarding and alighting, this study analysed the spatial and temporal variation in passenger service time. This paper introduced a time-series clustering approach, which uses the DTW algorithm for similarity measures and clustering of train doors by their temporal passenger service pattern. The analysis resulted in four distinct shapes of the passenger service rate curve from the north platform. For validation, a multivariate regression analysis was conducted, which included categorical dummy variables of the temporal service pattern cluster. Statistical tests proved the significance of the cluster variable. The validation results suggest that the spatial and temporal variation in passenger service time must be accounted for to accurately describe the relationship between passenger movements and the length of passenger service time.

This study will contribute to the literature by providing a new technical approach to analyse and predict levels of interference and its effect on the passenger service time. Interference can be minimised by distributing passengers along the platform and among train cars. Real-time passenger information may be an effective measure to guide passengers to less crowded train cars. Recently, Ahn et al. (2016) tested a strategy to provide the occupancy level of arriving train cars to passengers in order to assist their decision making for positioning and choice of a train car. The proposed approach may contribute to more sophisticated passenger information systems by deriving the optimal number of passengers to allocate to each designated door location to minimise interference and overall train dwell times.

The presented passenger service time model performed reasonably well using only four explanatory variables. The cluster variable is a constant, which can be defined through an off-line analysis of the historical passenger service time data. Train weighing scales are already in operation in many countries. The number of boarding passengers may be estimated through automatic fare collection information (Chan, 2007; Sun and Xu, 2012) or video surveillance (Liu et al., 2005; Zhang and Chen, 2007; Conte et al., 2010; Xu et al., 2010). Although more challenging, prediction of alighting passengers in reasonable accuracy will enable forecasting passenger service time and dwell time on the real-time basis.

8. References

Aach, J. and Church, G.M., 2001. Aligning gene expression time series with time warping algorithms. *Bioinformatics*, *17*(6), pp.495-508.

Ahn, S., Kim, J., Bekti, A., Cheng, L.C., Clark, E., Robertson, M., and Salita, R., 2016. Real-time Information System for Spreading Rail Passengers across Train Carriages: Agent-based Simulation Study, *Australasian Transport Research Forum 2016 Proceedings, Melbourne, Australia*

Berndt, D.J. and Clifford, J., 1994, July. Using dynamic time warping to find patterns in time series. In *KDD workshop* (Vol. 10, No. 16, pp. 359-370).

Caiani, E.G., Porta, A., Baselli, G., Turiel, M., Muzzupappa, S., Pieruzzi, F., Crema, C., Malliani, A. and Cerutti, S., 1998, September. Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume. In *Computers in Cardiology 1998* (pp. 73-76). IEEE.

Chan, J., 2007. *Rail transit OD matrix estimation and journey time reliability metrics using automated fare data* (Doctoral dissertation, Massachusetts Institute of Technology).

Chu, S., Keogh, E., Hart, D. and Pazzani, M., 2002, April. Iterative deepening dynamic time warping for time series. In *Proceedings of the 2002 SIAM International Conference on Data Mining* (pp. 195-212). Society for Industrial and Applied Mathematics.

Conte, D., Foggia, P., Percannella, G., Tufano, F. and Vento, M., 2010. A method for counting moving people in video surveillance videos. *EURASIP Journal on Advances in Signal Processing*, *2010*(1), p.231-240.

Coxon, S., Sarvi, M., Napper, R., and De Bono, A., 2013. An Innovative Approach to Metropolitan Train Carriage Interior Configuration: To Improve Boarding, Alighting, Dwell Time Stability and Passenger Experience, *Proceedings of the 13th World Conference on Transport Research, Rio de Janeiro, Brazil*

Douglas, N., 2012. Modelling Train & Station Demand & Capacity. *Report to Transport for NSW*.

Gavrila, D.M. and Davis, L.S., 1995. Towards 3-d model-based tracking and recognition of human movement: a multi-view approach. In *International workshop on automatic face-and gesture-recognition* (pp. 272-277).

Gollmer, K. and Posten, C., 1995. Detection of distorted pattern using dynamic time warping algorithm and application for supervision of bioprocesses. *On-line fault detection and supervision in chemical process industries*.

Harris, N.G. and Ramsey, J.B.H., 1994. Assessing the effects of railway infrastructure failure. *Journal of the Operational Research Society*, *45*(6), pp.635-640.

Harris, N.G., 2006. Train boarding and alighting rates at high passenger loads. *Journal of advanced transportation*, *40*(3), pp.249-263.

Jong, J.C. and Chang, E.F., 2011. Investigation and estimation of train dwell time for timetable planning. In *Proceedings of 9th World Congress on Railway Research*.

Kecman, P. and Goverde, R.M., 2013. *An online railway traffic prediction model*. International Association of Railway Operations Research (IAROR).

Keogh, E., Chakrabarti, K., Pazzani, M. and Mehrotra, S., 2001. Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and information Systems*, *3*(3), pp.263-286.

Keogh, E. and Kasetty, S., 2002. On the need for time series data mining benchmarks: a survey and empirical demonstration. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 102-111). ACM.

Keogh, E. and Ratanamahatana, C.A., 2005. Exact indexing of dynamic time warping. *Knowledge and Information Systems*, *3*(7), pp.358-386.

Kim, H., Kwon, S., Wu, S.K. and Sohn, K., 2014. Why do passengers choose a specific car of a metro train during the morning peak hours?. *Transportation research part A: policy and practice*, *61*, pp.249-258.

KOSIS., 2013. National Survey on Physical Activity [Online]. Ministry of Culture, Sports and Tourism. Available:

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1020215&conn_path=I2.

Kovacs-Vajna, Z.M., 2000. A fingerprint verification system based on triangular matching and dynamic time warping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), pp.1266-1276.

Krstanoski, N., 2014. Modelling passenger distribution on metro station platform. *International Journal for Traffic & Transport Engineering*, *4*(4).

Lam, W.H., Cheung, C.Y. and Poon, Y.F., 1998. A study of train dwelling time at the Hong Kong mass transit railway system. *Journal of advanced Transportation*, *32*(3), pp.285-295.

Li, D., Daamen, W. and Goverde, R.M., 2016. Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station. *Journal of Advanced Transportation*, *50*(5), pp.877-896.

Lin, T.M. and Wilson, N.H., 1992. Dwell time relationships for light rail systems. *Transportation Research Record*, (1361).

Liu, X., Tu, P.H., Rittscher, J., Perera, A. and Krahnstoever, N., 2005, September. Detecting and counting people in surveillance applications. In *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on* (pp. 306-311). IEEE.

Liu, Z., Li, D. and Wang, X., 2016. Passenger Distribution and Waiting Position Selection Model on Metro Station Platform.

Munich, M.E. and Perona, P., 1999. Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on* (Vol. 1, pp. 108-115). IEEE.

Parkinson, T. and Fisher, I., 1996. *Rail transit capacity* (Vol. 13). Transportation Research Board.

Puong, A., 2000. Dwell time model and analysis for the MBTA red line. *Massachusetts Institute of Technology Research Memo*.

Rabiner, L.R. and Juang, B.H., 1993. Fundamentals of speech recognition.

Sarle, W. S., 1983. *Cubic clustering criterion*. SAS Institute Incorporated.

Schmill, M.D., Oates, T. and Cohen, P.R., 1999. Learned models for continuous planning. In *AISTATS*.

Sun, Y. and Xu, R., 2012. Rail transit travel time reliability and estimation of passenger route choice behavior: Analysis using automatic fare collection data. *Transportation Research Record: Journal of the Transportation Research Board*, (2275), pp.58-67.

Szplett, D. and Wirasinghe, S.C., 1984a. An investigation of passenger interchange and train standing time at LRT stations:(i) alighting, boarding and platform distribution of passengers. *Journal of advanced transportation*, *18*(1), pp.1-12.

Szplett, D. and Wirasinghe, S.C., 1984b. An investigation of passenger interchange and train standing time at LRT stations:(ii) estimation of standing time. *Journal of advanced transportation*, *18*(1), pp.13-24.

Weston, J.G., 1989. Train service model–technical guide. *London underground operational research note*, *89*, p.18.

Wiggenraad, P.B.L., 2001. Alighting and boarding times of passengers at Dutch railway stations. *TRAIL Research School, Delft*.

Xu, H., Lv, P. and Meng, L., 2010, June. A people counting system based on head-shoulder detection and tracking in surveillance video. In *Computer Design and Applications (ICCDA), 2010 International Conference on* (Vol. 1, pp. V1-394). IEEE.

Zhang, E. and Chen, F., 2007. A fast and robust people counting method in video surveillance. In *Computational Intelligence and Security, 2007 International Conference on* (pp. 339-343). IEEE.