Transportation
https://doi.org/10.1007/s11116-024-10570-w

Investigating passenger behaviour on the metro platform with Wi-Fi location tracking data: a case study of Singapore

Michelle Cheung[3] · **Yan Cheng**[1,2] · **Taku Fujiyama**[3]

Accepted: 27 November 2024

Abstract

Utilising the existing infrastructure in railway transit to tackle overcrowding requires more understanding of how people use spaces at stations. This study investigated passenger behaviour while waiting for a train on the platform using the data of the Wi-Fi location tracking systems. The trajectories of 129,354 devices were observed in two weeks at two MRT Circle Line stations in Singapore, which have the escalator/stair landings in different positions. A data cleaning process was proposed to overcome the drawbacks of Wi-Fi-based position data. A decomposition method was further developed to separate the walking and staying phases based on data processing. The boarding passengers' on-platform behaviour was analysed from four aspects: the number of staying phases, the location distributions of different kinds of stays, the location distribution of in-between stays by hour and duration, and the distance and walking speed of the first walking phase. Our results suggested that many passengers (44% and 37% of passengers at the two case study stations) had multiple staying phases, meaning that they did not go directly to their final boarding points after coming to the platform but rather made stops or walkarounds before coming to boarding points. The distributions of locations of the last and in-between stays were significantly different and may influenced by the width, length and layout (such as landing locations) of stations. In addition, the walking speeds of passengers observed on the metro platform were slower than those observed on the streets. These findings indicated that some commonly used assumptions in most simulation models are not true according to the empirical observation. The obtained knowledge would deepen the understanding of the passengers' on-platform behaviour and thus provide implications for designing railway stations and planning station operations.

Keywords Passenger movement · Metro platform · Walking and staying phases · Wi-Fi location tracking data · Passenger distribution

Extended author information available on the last page of the article

Introduction

Context

Globally, many existing metros face ridership increases in recent years. Overcrowding on public transport becomes a more serious issue due to the limited capacity of infrastructure (Börjesson and Rubensson 2019; UITP 2018). Although the COVID-19 pandemic observed in the last couple of years has caused demand drops (Department for Transport 2022), the recent ridership update shows most urban rail transit systems have recovered gradually after the pandemic. In some parts of Asia, South America and Africa, it recovered relatively quickly and even surpassed pre-pandemic levels (SLOCAT 2023). In Singapore, public transport ridership hit 93.5% of pre-pandemic levels in 2023, and there were 3.45 million daily MRT and LRT rides in 2023, with a 17.6 per cent increase from the number of rides in 2022 (Loi 2024). Additionally, rail and public transport would play a key role in solving global societal challenges, such as decarbonisation (Department for Transport 2021).

Uncertainty in future demand and subsequent reduced investment in infrastructure would mean that more operational resilience is required for existing systems (Little 2020; Gkiotsalitis and Cats 2021), thus agile utilisation of existing infrastructure to tackle overcrowding would remain an important issue. The need for social distancing in transport vehicles during the COVID-19 pandemic led to attention to various soft measures that do not need infrastructure investment (Hörcher et al. 2021; Vickerman 2021). In such contexts, it is essential to understand how people use space on the train platform, and such an improved understanding would inform future station designs and demand management (Fang et al. 2019). This paper therefore aims to investigate the on-platform behaviour of passengers, in particular, passenger behaviour of walking and staying on the platform.

Literature review

Most of the existing research relevant to this work can fall into one (or more) of the following categories: passenger behaviour on the platform, and passenger distribution on the platform.

First, passenger movements and behaviours within train stations have been studied to better understand how to optimise human flows, improve journey time and maximise network capacity. There has been much research into passenger behaviour at the train platform interface (TPI) because any delay in passenger alighting and boarding would lead to train departure delays and affect train service performance. The passenger alighting/ boarding time and model development for it may be one of the most investigated topics (Harris 2006; Harris and Anderson 2007; Li et al. 2016; Lin and Wilson 1981; Palmqvist et al. 2020; Qu et al. 2023). While some researchers have run laboratory experiments to conduct a detailed investigation into the passenger flow rate (Daamen et al. 2008; Seriani et al. 2017; Zhang et al. 2022), there have been some studies that conducted observation at stations (Lee et al. 2007; Wiggenraad 2001) to investigate a range of factors, such as the queuing and crowding on the platform, that would affect alighting and boarding. In addition to these studies on passenger behaviour at the TPI, there have been studies that investigated passenger behaviour in other areas or situations, such as ramps (Fujiyama et al. 2015), seating behaviour on the platform (Schöttl et al. 2019) and behaviour in case of emergency (Shiwakoti et al. 2017). Furthermore, the on-platform behaviour of passengers includes the time dimension as well.

Waiting time within transport environments including transferring stations has been intensively researched because waiting time on platforms has many implications in transport planning and modelling, and some recent research using smartcard data allows detailed estimation of this (Ingvardson et al. 2018; Wahaballa et al. 2017).

The second group concerns passenger distribution on the platform. There have been studies that developed distribution models based on some assumptions and/or hypotheses (Bilde et al. 2022; Ding et al. 2021; Leurent and Liang 2022; Yang et al. 2018, 2017), but here we first examine empirical evidence, be it a questionnaire survey and/or video observation. Kim et al. (2014) conducted a questionnaire survey in the morning peak time at stations on a Seoul metro line and concluded that passengers chose their boarding cars in ways to minimise the walking distance at destination stations, whilst a Canadian study on passenger distribution on the platform can be modelled as multinomial probability distributions (Krstanoski 2014). There have been studies that suggest the distribution is not uniform but related to the (only) platform access point (Lam et al. 1999; Peftitsi et al. 2020; Wirasinghe and Szplett 1984), while others suggested concentration around particular carriages (Oliveira et al. 2019). Note that one limitation of these studies is that they either used questionnaires or videos from which the boarding points of passengers were extracted and hence did not record detailed movements associated with the time evolution. As an extension, some recent studies have turned attention to new technical solutions to guide passengers to other boarding locations (i.e., changing the boarder distribution) in order to improve train operation efficiency and customer satisfaction. The examined solutions include the provision of information about carriage occupancy and exit locations (Çapalar et al. 2018; Christoforou et al. 2017; Moncrieff 2015; Preston et al. 2017), lighting (Hughes et al. 2020), and alteration of train stopping positions (Van Den Heuvel 2016).

It is noteworthy that the emergence of new data sources, such as passive-collected big data, has offered new capabilities for behaviour analysis. Recent studies have exploited them, including train loading data (Fang et al. 2019; Peftitsi et al. 2020) and smart card data (Ingvardson et al. 2018; Leurent and Xie 2018; Zhu et al. 2017). However, one needs to be reminded that each new data source has its limitations: for example, train loading data does not distinguish passengers who just boarded at the focused stations and those who were already onboard from previous stations, and smart card data only infers passenger movements as it records the ticket barrier entry/exit times only. Thus, the on-platform behaviour of passengers over time and across the platform cannot be fully investigated by only using these two data sources.

In short, there has been research on passenger behaviour, but there is a lack of research on fundamental aspects of passenger movements on the platform, including how they wait for their trains. Accurate information on these will be useful for planning purposes and for the calibration of pedestrian simulation models which need detailed information on such aspects.

Study contributions

The contribution of our study to the existing body of knowledge is in the following points. First, Wi-Fi-based passenger position data, which is a new data source and has not been applied in the context of analysis of movements on train platforms will be used in this study.

Secondly, based on the analysis of the new data, empirical evidence on how passengers wait for trains on the platform will be provided.

The data used was obtained from the location estimation systems of Wi-Fi probe requests installed on the metro station platforms, which allowed analysis of passenger trajectories on the platform. To recognise the different behaviour phases of passengers, a decomposition method was adopted. Comparisons between the behaviours of passengers entering one station from different entry points and the ones at stations with different layouts were carried out to reveal the influence of facility locations on passengers' on-platform behaviour. This approach sheds light on the in-between stay behaviour of passengers, which significantly differs from the assumptions made by most simulation models, i.e., passengers who enter the platform would walk straight to preassigned boarding points and wait for trains. The distribution of in-between stays differs from the one of final boarding points and may influenced by the station layout. Furthermore, the walking speed of passengers is not as same as the one of pedestrians in normal open spaces. Such an approach will help clarify the nature of passengers' on-platform behaviour at a deeper understanding and enable transport designers to more precisely design the layout of stations and transport operators to better manage the demand within stations and especially on platforms.

Data

Characteristics of Wi-Fi location tracking data

As the data source, the data of Wireless@SG network was used, which provides free Wi-Fi services to the general public in various public spaces in Singapore including metro station platforms. The Wi-Fi probes are installed at intervals within the range of metro platforms to receive the requests. When Wi-Fi probe requests are sent by mobile devices, the system records information about devices including their media access control (MAC) addresses. There are various techniques to estimate the locations of devices (Zhu et al. 2020), and in the case of Wireless@SG system used for metro stations, its positional data has a 5-m radius margin of error as its positional accuracy according to the system operator. Its time resolution is seconds. A study in 2017 found that the proportion of the devices with randomised MAC addresses was less than 50% of the entire sample (Martin et al. 2017). At that time, mobile phone operation systems with enhanced MAC address randomisation (e.g., iOS14) were not commonly used. Whilst the size of the entire population is unknown, the dataset would allow investigation of the behaviour of passengers with observed devices. In addition, if a device is in sleep mode, the device does not send Wi-Fi probe requests (Chilipirea et al. 2018a). There have been studies that analysed Wi-Fi location from various viewpoints, including crowd movements (Chilipirea et al. 2018b; Gioia et al. 2019; Zhou et al. 2020), city-level pedestrian trajectory (Traunmueller et al. 2018; Zhou et al. 2020), and public transport loading (Wu et al. 2020). While the accuracy of Wi-Fi location tracking systems is deemed higher indoors than outdoors (Zhou et al. 2020), they have not been used for the investigation of passenger movements in stations or other railway environments.

Data acquisition

Passenger Wi-Fi data obtained from the platforms of selected stations during a selected period was acquired from the Land Transport Authority (LTA) of Singapore. The research ethics for this research was approved by University College London (UCL). The data acquisition period was from 30 September to 13 October 2019 (two weeks). Each data point contains the following information: timestamp, venue ID, pseudonymised MAC address, and coordinates of the estimated device location. LTA informed the research team that about 1 million unique MAC IDs were detected daily across the network around the time of the data collection. Estimating against the daily 3.5 million MRT trips (Tan 2019) and assuming that the average number of trips per passenger is 2 per day, the Wi-Fi location tracking system detected about 57% of passengers. The user characteristics (e.g., gender, age, etc.) of the Wi-Fi service were unknown to the authors.

Choice of stations

Bishan, Serangoon, Harbour Front, Buona Vista and Paya Lebar on the Circle Line (see Fig. 1), all of which are transfer stations to other lines, were chosen for initial data processing because LTA was undergoing an update and calibration of the Wi-Fi location estimation system and these stations had completed this by the time of the data acquisition. All the selected five stations have a rectangle island platform where the platform is surrounded by tracks in clockwise and anticlockwise directions (except Paya Lebar which has a siding between the two tracks), but the locations of the escalator/stair landings differ.

Fig. 1 Map of Singapore Circle Line

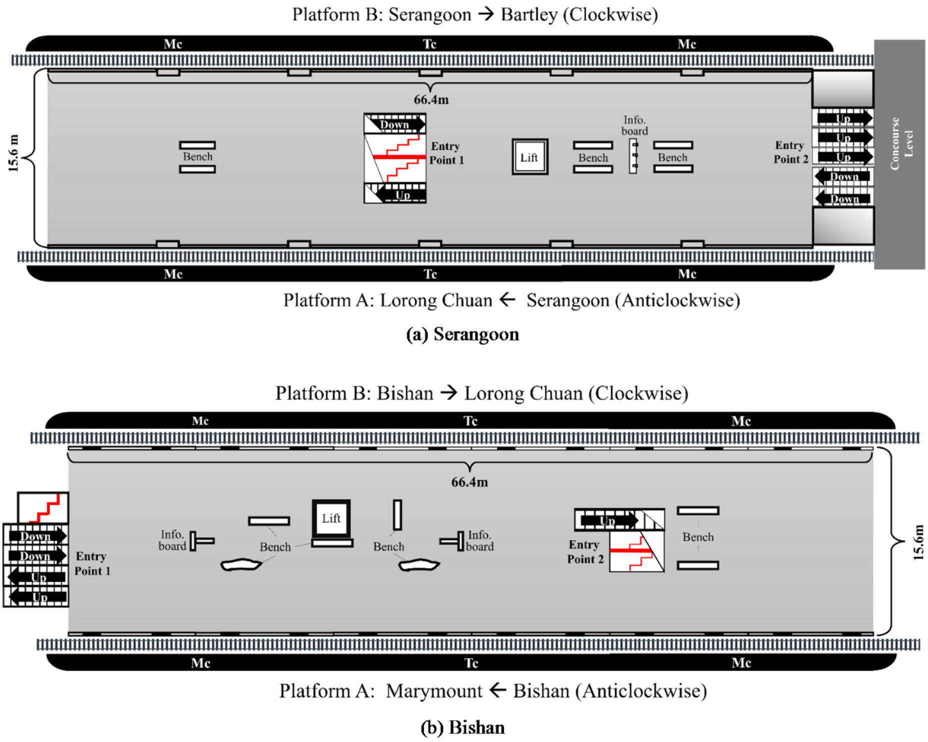


Fig. 2 A schematic representation of the platform layout of **a** Serangoon station and **b** Bishan station

Fig. 3 Photo of Serangoon station platform. *Source: Courtesy of LandTransportGuru.net website

This study compared the results of two stations: Serangoon with the landings in the middle and the east end of the platform and Bishan with landings in the middle and at the west end, as shown in Fig. 2. These two stations are near to each other with only one station (Lorong Chuan) between them. In our later analysis, the X-axis is set in parallel with the direction of the tracks, while the y-axis is at a right angle with the direction of the tracks. Figure 3 is a photo of Serangoon station platform. The average service frequencies of both stations ranged from 2-5 min in peak hours and 4-5 min during off-peak on weekdays, with 3.5-5 min on weekends. Platform edge doors are installed in both stations.

Methods

Data cleaning

The obtained dataset showed that on average, the position of each device was detected every 8.8 s, which may give sufficient frequency for trajectory identification of passengers. For extraction of data of only those who boarded trains, the datasets have two shortcomings: 1) location estimation has the aforementioned error range, and 2) the data include not only passengers who boarded a train at a focused station but also passengers who alighted from a train at the station and passengers who were staying inside a train or on the platform (without boarding/alighting) but has used station Wi-Fi systems. To overcome the two problems, a data cleaning process was designed as shown in Fig. 4.

For the first problem, the moving averaging was adopted. The corrected coordinate of a device at a time was an average of the detected coordinate of the time and those of the previous two data points of the same device. For the second problem, the data of devices that did not start around the position of the escalator/stair landing on the platform (i.e., with more than 5 m from the centre of the landing) were first excluded. Then the data whose Y coordinates (the dimension in a right angle with the track direction) did not move more than 2 m were removed. To ensure that the data of alighting passengers was excluded, only the data of devices which were observed at one of the other stations at a later time (i.e., within 1 h) of the same day were used. This step may also have helped to pick up the data of passengers who keep their mobile phones on. In addition, the data of devices that were observed to be on the platform for more than 10 min (twice the time of the off-peak frequency of the line) were filtered out because they could be people who did not intend to take the first available train (e.g., waiting for other people) or metro employees. Finally, the first two data points of the data of each device were removed as our initial data analysis showed that these data points were recorded when passengers may have been still on the escalator or the stairs.

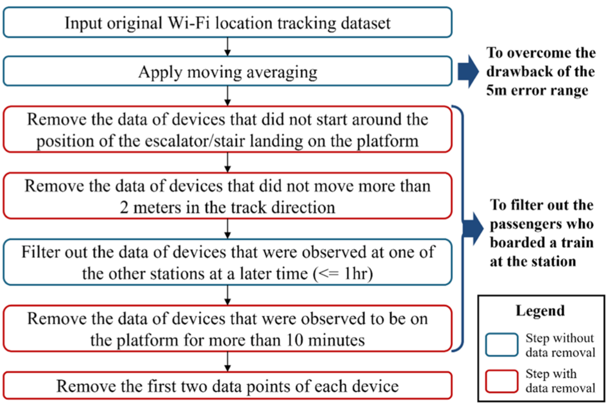


Fig. 4 Data cleaning process of Wi-Fi location tracking data

Decomposition of the walking and staying phases on the platform

Our preliminary analysis found that after entering the platform many passengers did not just continue walking until they boarded, that is, they may have walked towards a place where they waited to board a train, stayed/waited there until the next train came, and boarded the train. To further analyse this, it would be useful to segmentise passenger movements and their time-spending on the platform, into walking and staying phases. Here, we assume that, for any data point, if a device stays in the same position (within the 5 m, which is the error range in position estimation, from the observed point) for the next 20 s, the device was regarded as being in a staying phase for all the points within those 20 s. Phases other than staying phases were regarded as walking phases. In a typical case, a device entered the platform, continued moving, and entered a staying phase at a certain point, as shown in Fig. 5. However, in some cases, after the first staying phase, a device moved again and then entered another staying phase (and repeated this multiple times), which means that it moved away from the 5-m range from any point in the previous staying phase and entered a new staying phase. For each device, its last-observed location in the track direction on the platform was defined as the location of the last stay, i.e. boarding point.

Boarding passengers' on-platform behaviour analysis

The analysis of boarding passengers' on-platform behaviour included four aspects. Firstly, after decomposing the passengers' walking and staying phases, the number of staying phases was investigated to test the assumption of most simulation models that once passengers arrive on the platforms, they will directly walk to their final boarding points and stay

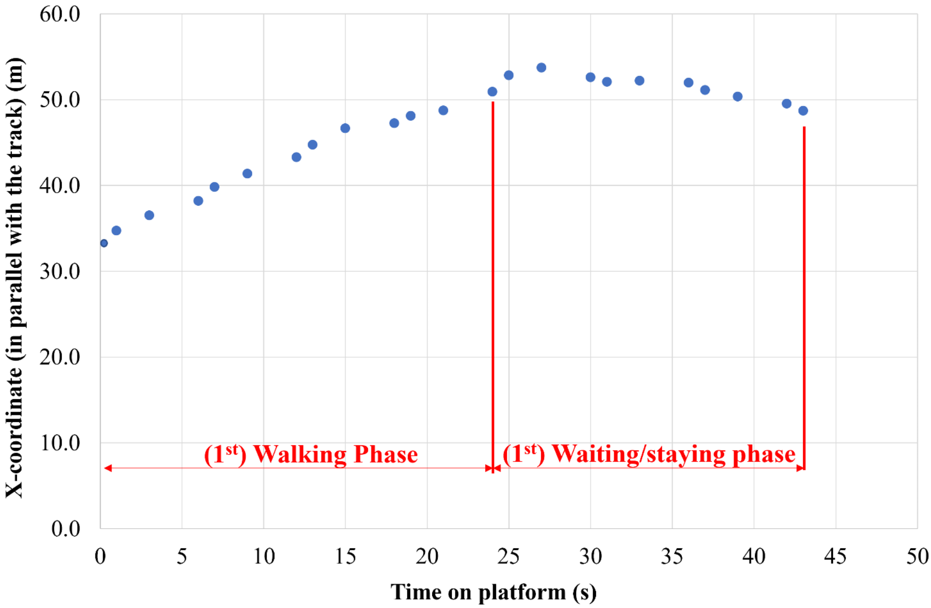


Fig. 5 Decomposition into walking and staying phases. *Please note that for illustration reasons, X coordinates were used as the vertical axis of the figure, but for the decision of whether a device stays within 5 m or not, both XY coordinates were considered

there waiting for a train. Secondly, the distributions of locations of different kinds of stays (i.e., the last stays and in-between stays), at stations with different layouts, and between passengers who arrived on the platforms from diverse entry points were compared by using the Kolmogorov–Smirnov Test (K-S test) to see whether significant differences existed in their shape. To further understand the behaviour of passengers with more than one staying phase, the third part focuses on the location distribution of in-between stays by hour and duration and how the distribution was influenced by station layout. Finally, to investigate the behaviour of passengers with multiple walking phases, we look into the first walking phase in detail. The ratio of distance along the track travelled in the first walking phase, to the distance along the track between the last observed point and the escalator/stair landing can be calculated according to Eq. (1).

$$R = \frac{x_1 - x_s}{x_L - x_s} \tag{1}$$

where x_s is the x-coordinate of the edge of the escalator/stair landing, x_1 is the x-coordinate of the first point of the first staying phase, and x_L is the x-coordinate of the last observed point of the device. For example, if a person walked 5 m along the track in the first walking phase but this person's last observed point shows a walking distance (along the track) of 10 m, the ratio is calculated as 5/10=0.5. If a person walked away from the escalator landing in the first walking phase but then later came back towards the landing, the ratio would be more than one. A negative ratio of a passenger means that that person walked to a certain place in the first walking phase, but then moved in the opposite direction in the next phases. A distance along the track was calculated from the difference between the x-coordinates of two points. Using the data of the first walking phase, a walking speed is calculated according to Eqs. (2) and (3).

$$v_t = \frac{\sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}}{t_{t+1} - t_t} \tag{2}$$

$$V = \sum_t \frac{t_{t+1} - t_t}{T} \cdot v_t \tag{3}$$

where V is the weighted average speed (m/s), x_t, y_t and v_t mean the x-coordinate, y-coordinate and speed at time t. T is the total duration in seconds a device was observed.

Results

The number of devices observed

Table 1 shows the number of observed devices that met our criteria detailed in the 'Method' section and hence were used for the analysis. In total, 129,354 devices were observed. Note that the same device can be regarded as different devices as it meets time separation criteria (i.e., 1 h). It shows one direction had more observations than the other and one hour in peak

Table 1 The number of observed devices that met the study criteria and the proportion in different time periods

Station	Direction		Mon	Tue	Wed	Thu	Fri	Sat	Sun
Serangoon	Anti-clockwise	Total	7643	8192	8306	7998	7690	6061	5013
		AM peak (%)	34.58	36.60	35.29	35.46	33.67	21.17	18.67
		PM peak (%)	14.38	13.90	14.70	15.13	15.06	16.38	16.84
		Off-peak (%)	51.04	49.50	50.01	49.41	51.27	62.45	64.49
	Clockwise	Total	3586	3808	3838	3886	3831	3882	3443
		AM peak (%)	41.41	43.36	41.48	41.89	37.54	19.37	23.76
		PM peak (%)	12.74	11.79	12.85	13.66	14.02	18.01	14.12
		Off-peak (%)	45.84	44.85	45.67	44.44	48.45	62.62	62.13
Bishan	Anti-clockwise	Total	2110	2226	2034	2179	2157	1309	1189
		AM peak (%)	47.54	46.09	44.44	46.08	41.40	26.74	19.85
		PM peak (%)	11.47	13.66	13.86	11.52	14.09	14.59	15.90
		Off-peak (%)	41.00	40.25	41.69	42.40	44.51	58.67	64.26
	Clockwise	Total	5484	5621	5523	5773	5849	5493	5230
		AM peak (%)	27.04	26.58	24.73	25.69	24.19	19.15	18.76
		PM peak (%)	18.85	21.40	19.30	18.34	20.21	18.84	18.91
		Off-peak (%)	54.10	52.02	55.97	55.97	55.60	62.01	62.33
	Grand total								129,354

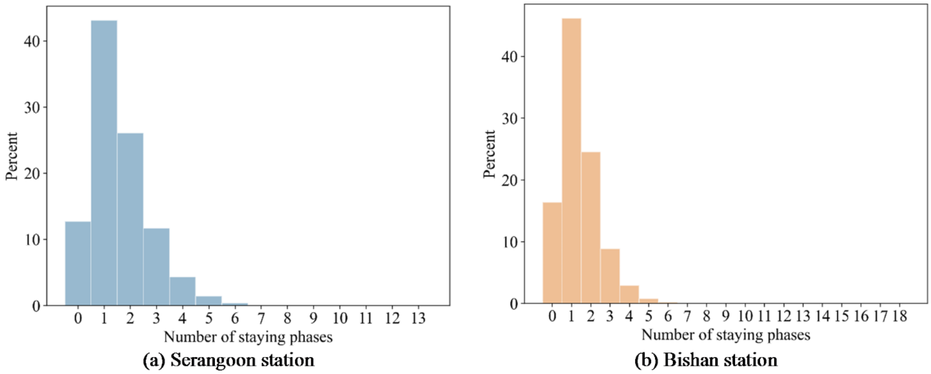


Fig. 6 Percentages of passenger numbers by their number of staying phases calculated

hours (7:00-9:59 and 17:00-19:59) has more observations than one hour during the off-peak (12 h). These may reflect the demand patterns of the stations.

Number of staying phases

According to the decomposition method in the 'Method' section, the percentages of passenger numbers by their number of staying phases calculated for each station can be shown in Fig. 6. Note that if a passenger did not have any staying phase within his/her data observed (e.g., he/she just came to a platform and jumped onto a train), then the passenger is counted with the number of staying phases = 0. Having one staying phase may mean that a passenger came to the platform, walked to a point where he/she stayed for a train, waited there for the next train, and then boarded. An interesting finding was that at Serangoon, 44.16% of passengers had two or more staying phases, while at Bishan 37.44%. These people may have walked to a certain point, stayed there, moved again to a different place, and waited in that new place until trains arrived.

According to the Kolmogorov-Smirnov test, the number of staying phases does not follow Normal, Poisson or any other typical distributions. Therefore, the Mann-Whitney U and Kruskal-Wallis tests were adopted to compare the means of the number of staying phases by different factors. Table 2 shows there were statistically significant differences between different stations ($p=0.000$), entry points ($p=0.000$), day types ($p=0.000$) and time periods ($p=0.000$). The mean number of staying phases of passengers in both stations was more than 1, with the passengers in Serangoon station moving more than the ones in Bishan station. This implies that passengers made moves after the first staying phases, which corresponds to the phenomenon Fig. 6 suggests. The difference between stations may be due to their station layouts. It was found that passengers arriving on the platforms from various entry points may have different walking and staying behaviour. Passengers with their entry points at the end of the stations (entry point 2 at Serangoon station and entry point 1 at Bishan station) had more moves on average than those entering the platform from the middle. The travelling direction of passengers showed a significant influence on their walking and staying behaviour in Serangoon station but not Bishan station.

Besides the factors related to the station layouts, the time-related factors also contribute to the difference in passengers' on-platform behaviour. The average number of staying phases during weekdays was 1.45, and at weekends the number was higher. Also, passengers who travelled during off-peak time may move more times than the ones during the AM and PM peaks. This could be attributed to the higher proportion of commuters who are more familiar with the station layout during peak hours on weekdays. During the off-peak period on weekdays and all periods on weekends, more passengers may travel for purposes other than commuting (such as for leisure, going to the hospital, business, etc.), so their destina-

Table 2 Mann-Whitney U and Kruskal-Wallis tests for the number of staying phases

Variable	Value	Mean	Std. deviation	Mann-Whitney U test statistic	Kruskal-Wallis test statistic	p-value
Station	Serangoon	1.59	1.15	1.83×10^9	-	0.000***
	Bishan	1.40	1.06			
Direction	Serangoon-Anticlockwise	1.53	1.11	7.22×10^8	-	0.000***
	Serangoon-Clockwise	1.71	1.23			
	Bishan-Anticlockwise	1.38	1.04	2.59×10^8	-	0.239
	Bishan-Clockwise	1.40	1.07			
Entry point	Serangoon-1	1.55	1.14	5.70×10^8	-	0.000***
	Serangoon-2	1.72	1.19			
	Bishan-1	1.48	1.10	1.22×10^8	-	0.000***
	Bishan-2	1.39	1.06			
Day type	Weekday	1.45	1.08	1.73×10^9	-	0.000***
	Weekend	1.71	1.22			
Time period	AM peak	1.40	1.04	-	1961.956	0.000***
	PM peak	1.30	0.98			
	Off-peak	1.64	1.19			

***Significant at $\alpha=0.01$

tion arrival time was more flexible and they were less familiar with the stations than the commuters. In addition, the train frequency during those periods was usually lower than the one in the peak periods on weekdays. Therefore, they may go to the information board or walk around to find out the best places to stay.

Locations of the last and in-between stays

To analyse where stays took place and whether that correlated with station designs, we divided passengers' stays into two groups: the last stay, also known as the boarding point, and the in-between stay. Figure 7 show the distributions of both kinds of stays in the track direction on the platforms at both stations during the two-week observation period. Each location segment covers 5 m longitudinally in the track direction (and laterally the whole platform width) and the numbers are distances (in meters) from the west end of the platform by comparing with the platform layout in Fig. 2. Numbers in clockwise (and anticlockwise)

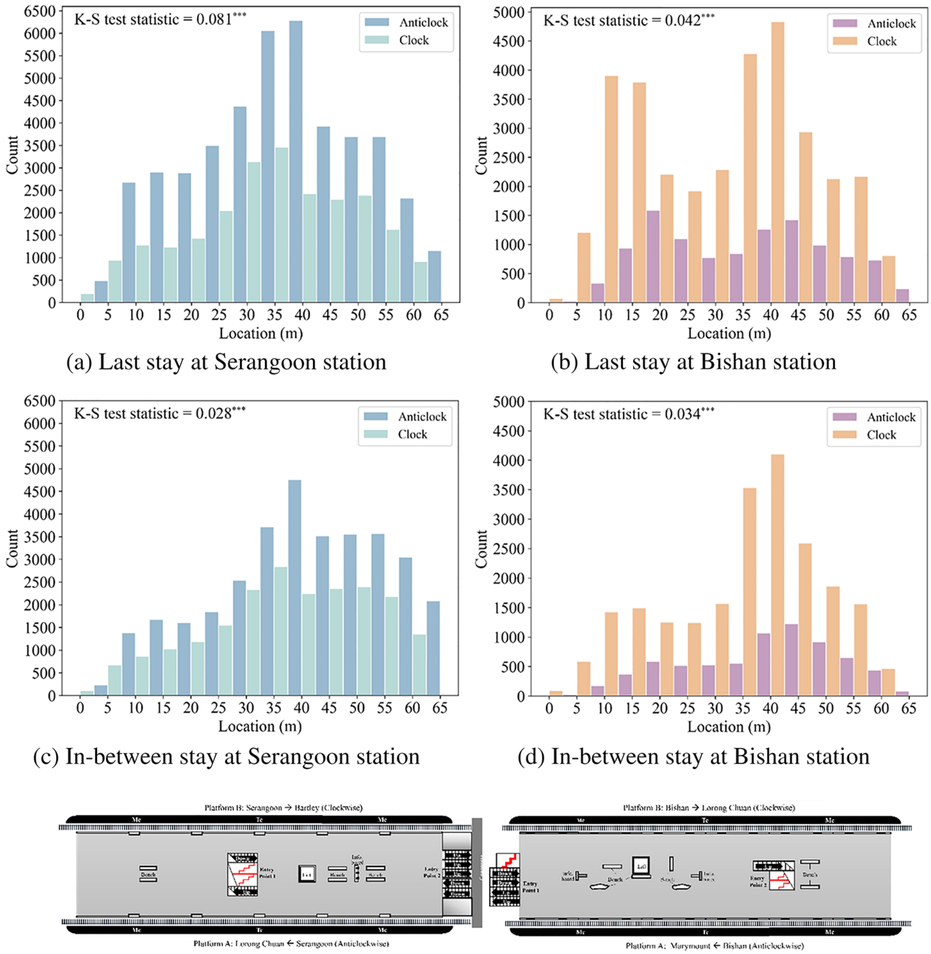


Fig. 7 Numbers of observed devices at their last and in-between stay points for each location segment by direction

direction mean the number of devices of passengers who travelled on clockwise (anticlockwise) trains. For example, in Fig. 7a, around 2600 devices were observed in the sector from 5 to 10 m in the track direction from the west end of the Serangoon station platform in the anticlockwise direction.

The histograms show while the counts (i.e., the number of devices observed) were different between the clockwise and anticlockwise directions for the last stays and in-between stays, their distributions seem similar between these directions. However, Kolmogorov-Smirnov test results revealed that the distributions of both kinds of stays were statistically different in the clockwise and anticlockwise directions. Nevertheless, both directions show concentration in similar areas. For the last stays, the distributions concentrated around the location 30-40 m at Serangoon station and 10-20 m and 35-45 m at Bishan station. It was interesting to find that one peak of Bishan station's distribution was in the region of 10-20 m, which suggests that people may walk along the platform anyway (not just wanting to minimise the walking distance at the boarding station).

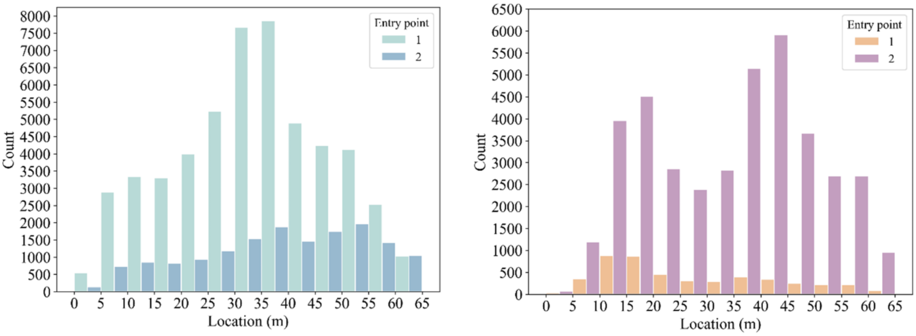
By comparing the distributions of the last stays in Fig. 7a and b and the ones of in-between stays in Fig. 7c and d at each station, it was easy to find out that the distributions of the locations of both kinds of stays were different. The K-S test results confirmed this with the statistic of Serangoon station equalling 0.119 and the one of Bishan station equalling 0.123, both p values less than 0.001. Specifically speaking, passengers tend to have their in-between stays on the east (or right) side of Serangoon station, while passengers prefer not to have their in-between stays on the west (or left) side of Bishan station.

To examine whether station layouts may affect the distributions, the distributions of the last and in-between stays of passengers arriving on the platforms from different entry points were further investigated. Figure 8 suggests there were significant differences in the distributions of both kinds of stays between the entry points. The final distributions of both stations shown in Fig. 7 were the accumulation of weighted distributions of passengers who entered from two entry points. Because the numbers of passengers utilizing the staircases and escalators were different, one of the distributions may dominate the total distribution of one station. For example, entry point 2 is much more used than entry point 1 so the final distributions of Bishan look very similar to the one of entry stair 2.

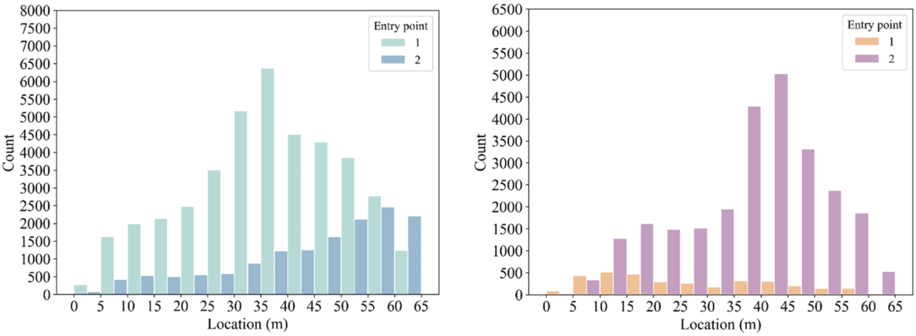
Location distributions of in-between stays by hour and duration

Figure 9 shows how locations of in-between stays are distributed along with the platform (track direction) within each time sector at both stations. For example, for the sector of 16:00-16:59, around 16% of the observed in-between stays were observed around the 35 m location at Bishan station. It was found that although there were differences in the distributions of stays between diverse entry points, all the distributions of the locations showed consistency over time. The in-between stays are more concentrated near the entry points. However, compared with the distributions of entry points in the middle of the platforms (i.e., entry point 1 at Serangoon station and entry point 2 at Bishan station), the ones of entry points at the end of the platforms (i.e., entry point 2 at Serangoon station and entry point 1 at Bishan station) were more dispersed.

We also divided observed in-between stays by their duration (into multipliers of 10 s) because for example, stays that lasted 30 s only may be different from those which lasted 180 s. Figure 10 shows the heatmaps divided by the duration sector. The longer stays tended



(a) Last stay at Serangoon station (b) Last stay at Bishan station



(c) In-between stay at Serangoon station (d) In-between stay at Bishan station

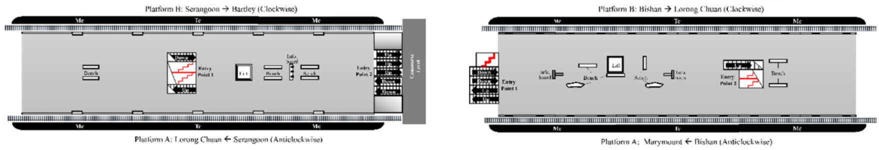


Fig. 8 Numbers of observed devices at their last and in-between stay points for each location segment by entry point

to be concentrated around the area with both bench and information board rather than other areas at both stations. The area with only benches available did not show a similar concentration. In addition, the passengers arriving from the entry points at the end of the platforms were more easily to have their longer stay at the closest bench and information board area.

First walking phase and its walking speed

Because the ratios of the distance (in x dimension) of the first walking phase to that of the total observed period, calculated by Eq. (1), may have very large absolute values due to a very short distance between the last observed point and the entry point, so 90 percentiles of the absolute ratios were used to filter out the majority of the ratios. Figure 11 shows the shapes of the violin plots of the ratios of the first walking phases to the total distance between boarding points and landings were different between the entry points in the middle and the ones at the end of the platforms. While the distributions of the entry points at the end of platforms during the AM peak and PM peak were similar, more passengers using these

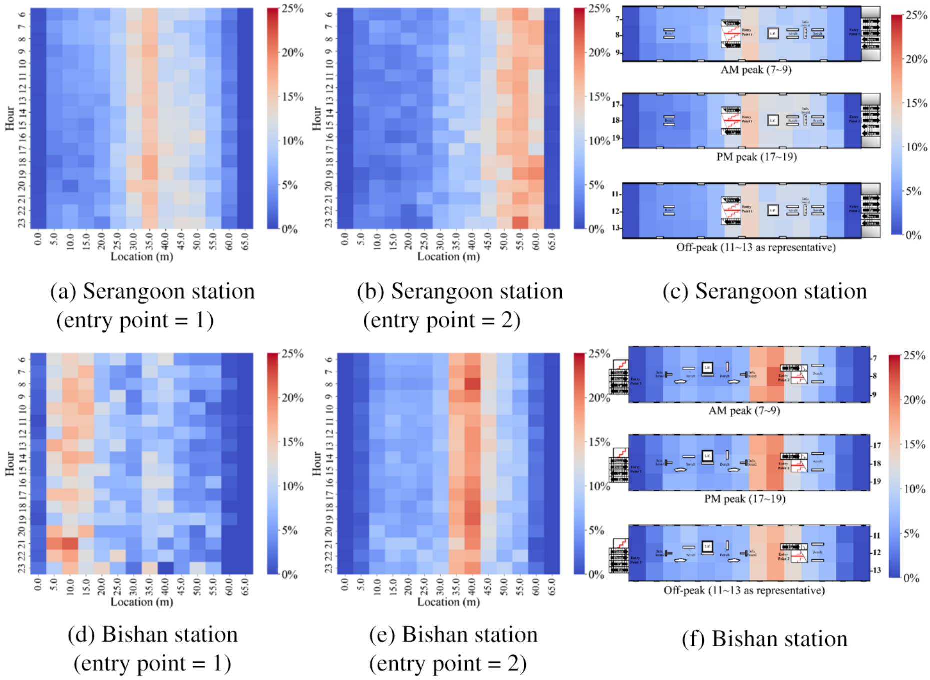


Fig. 9 Stay location heatmaps by hour and entry point

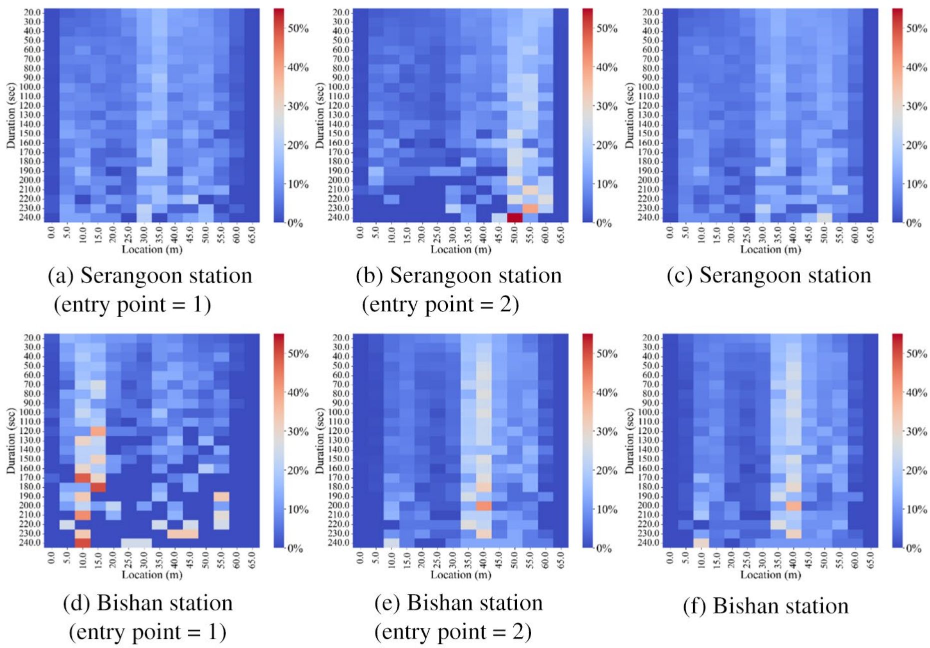
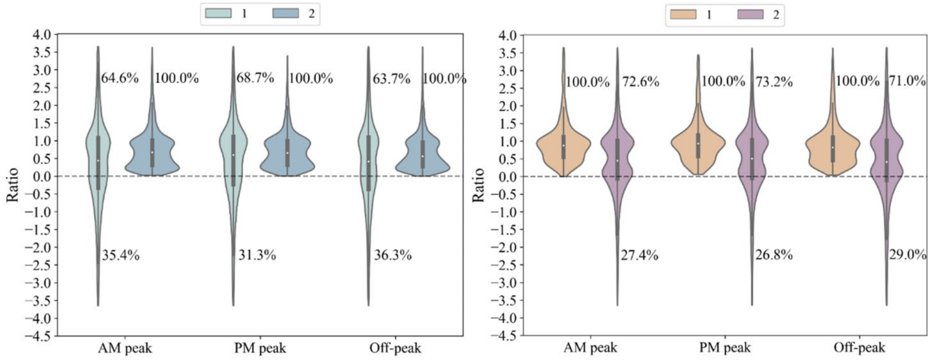
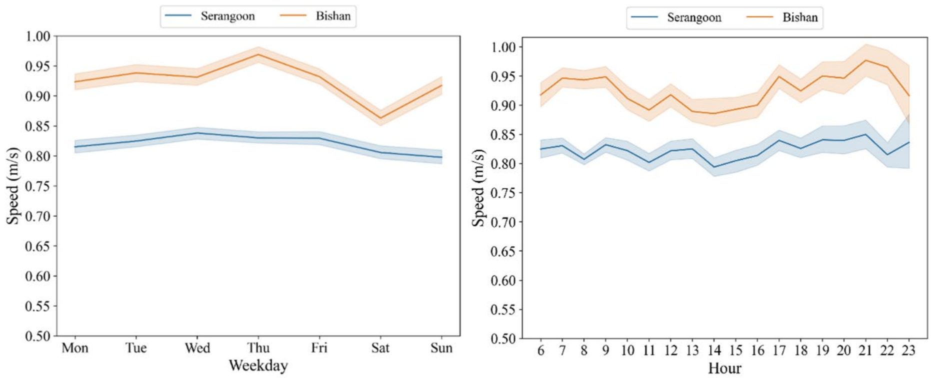


Fig. 10 Stay location heatmaps by duration and entry point



(a) Serangoon station (b) Bishan station

Fig. 11 Ratio of the walking distance (in X-dimension) in the first walking phase by time period



(a) By day (b) By hour

Fig. 12 Walking speeds of the first walking phases at both stations

entry points during the off-peak period walked relatively shorter (about 25% of the distance of the last observed point from the entry point) in their first walking phases. It also suggests that about one-third of passengers arriving from the entry points located in the middle of the platforms walked to the opposite side of the entry point compared with the last observed point, which indicated that passengers who arrived on the platform from the middle may hesitate about which direction to go at the beginning of their walking, with the proportion of entry point 1 on Serangoon station's platform higher than the one of entry point 2 on Bishan station's platform.

Figure 12 shows walking speeds of the first walking phases at both stations by day and hour. Interestingly, both stations had walking speeds lower than 1 m/s, with Bishan having faster speeds (average 0.93 m/s) than Serangoon (0.82 m/s). This may be because walking distances at Bishan are longer than the ones at Serangoon and this walking for longer distances may have contributed to faster speeds; people can walk faster if they walk longer. On weekdays, passengers with a higher proportion of commuting and other fixed travel purposes walked faster than the ones travelling during weekends. Two stations showed similar patterns over one day. It suggests that the passengers travelling in the AM peak and PM peak preferred to walk at a higher speed than the ones travelling during the midday off-peak, However, probably due to the on-platform crowdedness, the walking speeds between 8:00–8:59 and 18:00–18:59 were lower than the other hours in peak periods.

Discussion

One important finding was that passengers often restart walking after an initial stay; 44% and 37% of observed passengers had two or more staying phases at Serangoon and Bishan stations respectively (Fig. 6). In fact, in city parks people (who are waiting for other friends) may not always stand still but make some little movements around. In their experiments using a mock-up platform, Küpper and Seyfried (2023) also found that with an increased waiting time, the area covered by pedestrians' head trajectories increases, as some participants start to walk around instead of waiting at a fixed position. In some metro environments where passengers may not be keenly seeking seats to sit on the train because of short rides, it would be reasonable to assume similar behaviour. The case study stations have wide (whose width is around 15 m) but shorter platforms (whose length is around 66 m), which may have made such little movements easy to occur. Besides, several potential reasons can explain the phenomenon of multiple stays during peak hours. Firstly, this may be because if there is a shorter queue, a passenger may move to that queue after joining another longer one. Secondly, passengers tend to avoid spaces just next to the platform entry place as their boarding positions (Fang et al. 2023), maybe because they do not want many new passengers coming to (or passing by). Therefore, some passengers may move to new places after finding out that their next train does not come soon. Lastly, obstructions to passengers' vision can affect passenger choice of boarding cars, because such obstructions can hinder passengers from finding less crowded areas on the platform to stand (Wu et al. 2010). Similarly, crowds of people (e.g. before the staircase and escalator) can also be obstacles to other passengers. After a crowd diminishes, some passengers may discover less crowded areas and subsequently make a move.

Our analysis of the locations of stays found that the distributions of the in-between stays shown in Fig. 7c and d were different from the distributions of final boarding points shown in Fig. 7a and b. Compared to the distribution of the last stay locations, Serangoon had a higher proportion of in-between stays in the locations from 40 to 65 m, while at Bishan they were from 30 to 60 m. At the aggregate level, the total distribution of the stays at a station was determined by two factors: 1) the distribution of each entry point and 2) the number of passengers using each entry point, as shown in Fig. 8. No matter the locations of the last stays or in-between stays, the shapes of the distributions by entry points on the platforms were significantly different. Its relationship with the distance between the stairs/escalators on the platform may need more investigation. It is noteworthy that if the number of passengers using each entry point changes because of the changes in the external environment (such as land use) and/or facility (such as new bus stops and shops), the distribution may vary.

It indicates that more passengers may stay a bit distance before the staircases/escalators in the middle of both stations, as they may need to find out which direction they should go, which was also suggested by Fig. 11. Fang et al. (2023) observed that in the choice of boarding place for the train, passengers tend to avoid spaces just outside the entrance to the platform maybe because they do not want to obstruct other passengers who arrive later. This may explain why very few passengers want to stay around the distance of 0 m at Bishan. The other potential reason why people tend not to stay on the west (or left) side of Bishan station is that the passengers entering the platform from the west end escalators/stairs may have a clearer view of the train directions given the wide platform and long distance from the west end landings to the middle landings, so they may have higher chances to direct to their first staying points (which may not be the boarding points) instead of walking around and searching for which side to go to.

In addition, the spaces where many stays took place include information boards and seating areas and are surrounded by the fronts of escalators/stairs and these would make the space attractive for staying. A relatively wide platform (around 16 m for the island platform at both stations) may have increased such attractiveness. Interestingly, there are seating spaces (around 10 m at Serangoon and 50 m at Bishan) where not many stays were observed partly because these spaces are under or facing the back side of staircases/escalators and may not be attractive for staying. Figure 10b and d show that some of the stays that lasted more than 170 s at Serangoon and around 120 s at Bishan may have taken around seating spaces, but the total number of such long stays was limited and hence was not reflected in Fig. 10c and f.

The ratio of the distance passengers walk in the first phase to the distance to the final boarding points was also different between entry points. The shapes of the distributions of entry points in the middle of both station platforms were similar, but the ones of the entry points at the end of station platforms varied. The distance from popular escalators/stairs to such popular staying places may have affected this ratio (Fig. 11). At Bishan, entry point 1 at the west end is farther away from the popular staying place on the east side.

It was found that the average walking speeds observed were around 0.9 m/s. The average walking speeds at Bishan station were faster than those of Serangoon, and one reason for this could be that the average walking distance is longer than Serangoon, suggesting that at Bishan passengers would walk faster as they (want to) walk longer. These walking speeds are relatively slower than ones observed on pedestrian crossings and other public spaces (Daamen and Hoogendoorn 2006; Montufar et al. 2007), which were in a range of 1.0-1.4 m/s. Specifically, according to previous studies on average walking speeds of pedestrians in Singapore, the average walking speeds is about 1.23 m/s on sidewalks along the main streets (Tanaboriboon et al. 1986) and 1.20 m/s in link-ways (Sun et al. 2021). Walking speed is greatly influenced by pedestrian environment and the characteristics of the surrounding crowd, and there is a big variation in the walking speed at various facilities, under different conditions and for various classes of pedestrians (Vanumu et al. 2017). There has been a study that suggests that pedestrians on different tasks walk less fast as usual suggesting that pedestrians on different tasks walk more slowly (Crowley et al. 2021), and finding boarding or stopping places could be such an additional task, resulting in reduced walking speeds.

Conclusions

Using data from wi-fi probe requests in Singapore, this study investigated the on-platform behaviour of passengers. Specifically, this research decomposed the passenger waiting behaviour into walking and staying phases. The research had the following findings.

- Passengers had multiple walking phases with stays in between, and the number of staying phases was influenced by the width, length and layout of stations.
- The distribution of the stay locations was an aggregation of various distributions of entry points. Both their locations and the number of passengers using these entry points may change the final distribution.
- Passengers concentrated near the entry points, but not just next to them, while the area with both bench and information board attracted more longer stays.
- Passengers arriving on the platform from the middle may hesitate about which side to go.
- Passengers' walking speeds were slower than those observed on the streets.

As a planner or operator of the station, a more even distribution is expected to avoid some point or area of crowding on the platform. These findings would have implications for crowd management and station designs. An assumption in crowd modelling is that boarding passengers would just walk straight into their preferred boarding points from escalator/stair landings at their normal speeds and wait there just before boarding. Such an assumption may be correct if joining a queue at a boarding point would provide some benefit (e.g., a better chance of getting a seat), but this may not be the case in certain cases. Firstly, if boarding passengers do not stay in the same places while waiting for trains, then the effective space on the platform (which is not occupied by stationary objects or passengers) could be larger than that calculated, so the width of platforms could be adjusted when designing a station platform. Secondly, given the forecast number of passengers and a better understanding of the shapes of the last and in-between stay locations of passengers arriving on the platforms from different entry points, planners and designers can improve the layout of entry and exit of a station. Last but not least, current platform designs rarely have considered such comfort of making little movements or the attractiveness of staying, which should be taken into account in the future design of station platforms. It was found that areas that have information boards or face the front of stairs/escalators attracted many stays, while areas under or facing the back of stairs/escalators as well as just next to the entrance landing to the platform did not. Besides, the findings can influence operational practice in metro systems. Some passengers who have reached their initially intended boarding point may be willing to move to different boarding positions, and hence be willing to respond to certain behaviour nudging, such as announcements suggesting less busy cars. Providing more real-time on-platform information has the potential to help reshape the distribution on the platform.

Wi-Fi data is a new data source capable of capturing the on-platform behaviour of metro passengers, but it still has some limitations. First, it can be used to estimate the relative levels of crowding but is unable to identify the absolute number of passengers. The 5-m error in its postitional data hinders it from being used to calculate the crowding level with high spatial resolution (e.g. 1 m$\times$1 m grid). It should be noted that nowadays, there are usually multiple types of cards and fare collection systems or tickets used in any metro systems and most types of data cannot cover all the passengers, estimating the absolute number of passengers is challenging and may need data fusion and new technologies. For example, computer vision observation could be a way forward, but occlusion often occurs in high density. The second limitation is that by using Wi-Fi data only, it is difficult to detect travel companions, which might influence passengers' behaviour. Wi-Fi location tracking data only generate when the probe requests are sent by mobile devices and it has a 5-m error, so it cannot compare inter-personal distances to identify social groups.

Further research can be conducted on the extent to which each platform design element (information board, seats, etc.) contributed to stays. With the help of additional datasets (such as survey data and passive data), the mechanism of passenger behaviour can be further investigated, such as the influence of crowdedness on passenger behaviour and the interactions between individuals.

Acknowledgements This research was supported by the Fundamental Research Funds for the Central Universities (No. 22120240058) granted by Tongji University, which funded the open access fees.

Author contributions Conceptualisation (MC), data curation (MC), methodology (MC+YC+TF), software (YC+TF), formal analysis (MC+YC+TF), investigation (MC+YC+TF), writing-original draft (MC+TF), writing-review & editing (MC+YC+TF), visualisation (YC+TF), supervision (TF), funding acquisition (YC). All authors reviewed the results and approved the final version of the manuscript.

Data availability The data that support the findings of this study are available from Land Transport Authority but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

Declarations

Conflict of interest The authors declare no competing interests.

References

Bilde, B.A., Andersen, M.L., Harrod, S.: Social distance modeling on the Copenhagen, Denmark, Metro. J. Transp. Eng. Part A Syst. **148** (2022). https://doi.org/10.1061/JTEPBS.0000633

Börjesson, M., Rubensson, I.: Satisfaction with crowding and other attributes in public transport. Transp. Policy **79**, 213–222 (2019). https://doi.org/10.1016/j.tranpol.2019.05.010

Çapalar, J., Nemec, A., Zahradnik, C., Olaverri-Monreal, C.: Optimization of passenger distribution at metro stations through a guidance system. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 397–404. Springer Verlag (2018)

Chilipirea, C., Baratchi, M., Dobre, C., van Steen, M.: Identifying stops and moves in WiFi tracking data. Sensors **18**, 1–15 (2018a). https://doi.org/10.3390/s18114039

Chilipirea, C., Dobre, C., Baratchi, M., Van Steen, M.: Identifying movements in noisy crowd analytics data. In: Proc. - IEEE Int. Conf. Mob. Data Manag. 2018-June, pp. 161–166 (2018b). https://doi.org/10.1109/MDM.2018.00033

Christoforou, Z., Collet, P.A., Kabalan, B., Leurent, F., de Feraudy, A., Ali, A., Arakelian-Von Freeden, T.J., Li, Y.: Influencing longitudinal passenger distribution on railway platforms to shorten and regularize train dwell times. Transp. Res. Rec. **2648**, 117–125 (2017). https://doi.org/10.3141/2648-14

Crowley, P., Vuillerme, N., Samani, A., Madeleine, P.: The effects of walking speed and mobile phone use on the walking dynamics of young adults. Sci. Rep. **11** (2021). https://doi.org/10.1038/s41598-020-79584-5
Daamen, W., Lee, Y., Wiggenraad, P.: Boarding and alighting experiments: overview of setup and performance and some preliminary results. Transp. Res. Rec. J. Transp. Res. Board. **2042**, 71–81 (2008). https://doi.org/10.3141/2042-08
Daamen, W., Hoogendoorn, S.P.: Free speed distributions for pedestrian traffic. In: Proc. 85th Annu. Meet. Transp. Res. Board., pp. 13–25 (2006)
Department for Transport. Decarbonising Transport (2021)
Department for Transport. Transport use during the coronavirus (COVID-19) pandemic (2022). https://www.gov.uk/government/statistics/transport-use-during-the-coronavirus-covid-19-pandemic
Ding, H., Di, Y., Zheng, X., Liu, K., Zhang, W., Zheng, L.: Passenger arrival distribution model and riding guidance on an urban rail transit platform. Phys. A Stat. Mech. its Appl. **571** (2021). https://doi.org/10.1016/j.physa.2021.125847
Fang, J., Fujiyama, T., Wong, H.: Modelling passenger distribution on metro platforms based on passengers' choices for boarding cars. Transp. Plan. Technol. **42**, 442–458 (2019). https://doi.org/10.1080/03081060.2019.1609218
Fang, J., Wong, H., Fujiyama, T.: Modelling Distributions of Heterogeneous Boarders and Alighters on Metro Platforms. Public Transp. (Under review) (2023)
Fujiyama, T., Childs, C., Boampong, D., Tyler, N.: Investigating ramp gradients for humps on railway platforms. Proc. Inst. Civ. Eng. Munic. Eng. **168**, 150–160 (2015). https://doi.org/10.1680/muen.14.00011
Gioia, C., Sermi, F., Tarchi, D., Vespe, M.: On cleaning strategies for WiFi positioning to monitor dynamic crowds. Appl. Geomatics. **11**, 381–399 (2019). https://doi.org/10.1007/s12518-019-00260-z
Gkiotsalitis, K., Cats, O.: Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. Transp. Rev. **41**, 374–392 (2021). https://doi.org/10.1080/01441647.2020.1857886
Harris, N.G.: Train boarding and alighting rates at high passenger loads. J. Adv. Transp. **40**, 249–263 (2006). https://doi.org/10.1002/atr.5670400302
Harris, N.G., Anderson, R.J.: An international comparison of urban rail boarding and alighting rates. Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. **221**, 521–526 (2007). https://doi.org/10.1243/09544097JRRT115
Hörcher, D., Singh, R., Graham, D.J.: Social distancing in public transport: mobilising new technologies for demand management under the Covid-19 crisis. Transportation (2021). https://doi.org/10.1007/s11116-021-10192-6
Hughes, N., Ryan, B., Hallewell, M., Coad, N., Grant, A., Parrott, N., Roberts, S., Thompson, K.: Identifying new concepts for innovative lighting-based interventions to influence movement and behaviours in train stations. Light. Res. Technol. **52**, 976–990 (2020). https://doi.org/10.1177/1477153520904405
Ingvardson, J.B., Nielsen, O.A., Raveau, S., Nielsen, B.F.: Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: a smart card data analysis. Transp. Res. Part C Emerg. Technol. **90**, 292–306 (2018). https://doi.org/10.1016/j.trc.2018.03.006
Kim, H., Kwon, S., Wu, S.K., Sohn, K.: Why do passengers choose a specific car of a metro train during the morning peak hours? Transp. Res. Part A Policy Pract. **61**, 249–258 (2014). https://doi.org/10.1016/j.tra.2014.02.015
Krstanoski, N.: Modelling passenger distribution on metro station platform. Int. J. Traffic Transp. Eng. **4**, 456–465 (2014). https://doi.org/10.7708/ijtte.2014.4(4).08
Küpper, M., Seyfried, A.: Waiting in crowded places: influence of number of pedestrians, waiting time andobstacles. J. R. Soc. Interface. **20**, (2023). https://doi.org/10.1098/rsif.2023.0193
Lam, W.H.K., Cheung, C.-Y., Lam, C.F.: A study of crowding effects at the Hong Kong light rail transit stations. Transp. Res. Part A Policy Pract. **33**, 401–415 (1999). https://doi.org/10.1016/S0965-8564(98)00050-0
Lee, Y., Daamen, W., Wiggenraad, P.: Boarding and alighting behavior of public transport passengers. In: Transportation Research Board 86th Annual Meeting (2007)
Leurent, F., Liang, K.: How do individual walk lengths and speeds, together with alighting flow, determine the platform egress times of train users? J. Adv. Transp. **2022** (2022). https://doi.org/10.1155/2022/3633293
Leurent, F., Xie, X.: On individual repositioning distance along platform during train waiting. J. Adv. Transp. **2018** (2018). https://doi.org/10.1155/2018/4264528
Li, D., Daamen, W., Goverde, R.M.P.: Estimation of train dwell time at short stops based on track occupation event data: a study at a Dutch railway station. J. Adv. Transp. **50**, 877–896 (2016). https://doi.org/10.1002/atr.1380
Lin, T., Wilson, N.H.M.: Dwell time relationships for light rail systems. Transp. Res. Rec. **1361**, 287–295 (1981)
Little, A.D.: UITP: The Future of Mobility post-COVID (2020)

Loi, E.: Public transport ridership hit 93.5% of pre-pandemic levels in (2024). https://www.straitstimes.com/singapore/transport/public-transport-ridership-hit-935-of-pre-pandemic-levels-in-2023

Martin, J., Mayberry, T., Donahue, C., Foppe, L., Brown, L., Riggins, C., Rye, E.C., Brown, D.: A study of MAC address randomization in mobile devices and when it fails. Priv. Enhanc. Technol. **4**, 365–383 (2017). https://doi.org/10.1515/popets-2017-0054

Moncrieff, K.: Designing passenger information for dwell time to support thameslink high capacity infrastructure. In: 5th International Rail Human Factors Conference (2015)

Montufar, J., Arango, J., Porter, M., Nakagawa, S.: Pedestrians' normal walking speed and speed when crossing a street. Transp. Res. Rec. J. Transp. Res. Board **2002**, 90–97 (2007). https://doi.org/10.3141/2002-12

Oliveira, L.C., Fox, C., Birrell, S., Cain, R.: Analysing passengers' behaviours when boarding trains to improve rail infrastructure and technology. Robot. Comput. Integr. Manuf. **57**, 282–291 (2019). https://doi.org/10.1016/j.rcim.2018.12.008

Palmqvist, C.W., Tomii, N., Ochiai, Y.: Explaining dwell time delays with passenger counts for some commuter trains in Stockholm and Tokyo. J. Rail Transp. Plan. Manag. **14** (2020). https://doi.org/10.1016/j.jrtpm.2020.100189

Peftitsi, S., Jenelius, E., Cats, O.: Determinants of passengers' metro car choice revealed through automated data sources: a Stockholm case study. Transp. A Transp. Sci. **16**, 529–549 (2020). https://doi.org/10.1080/23249935.2020.1720040

Preston, J., Pritchard, J., Waterson, B.: Train overcrowding: Investigation of the provision of better information to mitigate the issues. Transp. Res. Rec. **2649**, 1–10 (2017). https://doi.org/10.3141/2649-01

Qu, Y., Liu, X., Wu, J., Wei, Y.: Modeling pedestrian behaviors of boarding and alighting dynamics in urban railway stations. Transp. A Transp. Sci. **19** (2023). https://doi.org/10.1080/23249935.2022.2035845

Schöttl, J., Seitz, M.J., Köster, G.: Investigating the randomness of passengers' seating behavior in suburban trains. Entropy **21** (2019). https://doi.org/10.3390/e21060600

Seriani, S., Fujiyama, T., Holloway, C.: Exploring the pedestrian level of interaction on platform conflict areas at metro stations by real-scale laboratory experiments. Transp. Plan. Technol. **40**, 100–118 (2017). https://doi.org/10.1080/03081060.2016.1238574

Shiwakoti, N., Tay, R., Stasinopoulos, P., Woolley, P.J.: Likely behaviours of passengers under emergency evacuation in train station. Saf. Sci. **91**, 40–48 (2017). https://doi.org/10.1016/j.ssci.2016.07.017

SLOCAT. Global Status Report on Transport, Climate and Sustainability, 3rd edn (2023)

Sun, S., Zhou, Q., Lal, S., Xu, H., Goh, K., Wong, Y.D.: Quantifying performance of sheltered link-way facility in Singapore using human-centric indicators. Int. J. Urban Sustain. Dev. **13**, 187–198 (2021). https://doi.org/10.1080/19463138.2020.1858422

Tan, C.: Bus and train ridership up, taxi rides down. https://www.straitstimes.com/singapore/transport/bus-and-train-ridership-up-taxi-rides-down (2019)

Tanaboriboon, Y., Hwa, S.S., Chor, C.H.: Pedestrian characteristics study in Singapore. J. Transp. Eng. **112**, 229–235 (1986). https://doi.org/10.1061/(ASCE)0733-947X(1986)112:3(229)

Traunmueller, M.W., Johnson, N., Malik, A., Kontokosta, C.E.: Digital footprints: using WiFi probe and locational data to analyze human mobility trajectories in cities. Comput. Environ. Urban Syst. **72**, 4–12 (2018). https://doi.org/10.1016/j.compenvurbsys.2018.07.006

UITP. New Urban Rail Infrastructure 2018 (2018)

Van Den Heuvel, J.: Field experiments with train stopping positions at Schiphol Airport train station in Amsterdam, Netherlands. Transp. Res. Rec. **2546**, 24–32 (2016). https://doi.org/10.3141/2546-04

Vanumu, L.D., Ramachandra Rao, K., Tiwari, G.: Fundamental diagrams of pedestrian flow characteristics: a review. Eur. Transp. Res. Rev. **9** (2017). https://doi.org/10.1007/s12544-017-0264-6

Vickerman, R.: Will Covid-19 put the public back in public transport? A UK perspective. Transp. Policy **103**, 95–102 (2021). https://doi.org/10.1016/j.tranpol.2021.01.005

Wahaballa, A.M., Kurauchi, F., Yamamoto, T., Schmöcker, J.D.: Estimation of platform waiting time distribution considering service reliability based on smart card data and performance reports. Transp. Res. Rec. **2652**, 30–38 (2017). https://doi.org/10.3141/2652-04

Wiggenraad, P.B.L.: Alighting and boarding times of passengers at Dutch railway stations. Delft (2001)

Wirasinghe, S.C., Szplett, D.: An investigation of passenger interchange and train standing time at LRT stations: (ii) estimation of standing time. J. Adv. Transp. 13–24 (1984). https://doi.org/10.1002/atr.5670180103

Wu, F., Yang, Z., Yuan, Y.: Waiting location choice of passengers in urban rail transit platform during the train stop. Urban Mass Transit. **13**, 52–56 (2010)

Wu, F.-J., Huang, Y., Doring, L., Althoff, S., Bitterschulte, K., Chai, K.Y., Mao, L., Grabarczyk, D., Kovacs, E.: PassengerFlows: a correlation-based passenger estimator in automated public transport. IEEE Trans. Netw. Sci. Eng. **7**, 2167–2181 (2020). https://doi.org/10.1109/TNSE.2020.2998536

Yang, X., Dong, H., Yao, X.: Passenger distribution modelling at the subway platform based on ant colony optimization algorithm. Simul. Model. Pract. Theory **77**, 228–244 (2017). https://doi.org/10.1016/j.simpat.2017.03.005

Yang, X., Yang, X., Wang, Z., Kang, Y.: A cost function approach to the prediction of passenger distribution at the subway platform. J. Adv. Transp. **2018** (2018). https://doi.org/10.1155/2018/5031940

Zhang, H., Xu, J., Jia, L., Shi, Y.: Modelling the walking behavior of pedestrians in the junction with chamfer zone of subway station. Phys. A Stat. Mech. Appl. **602** (2022). https://doi.org/10.1016/j.physa.2022.127656

Zhou, Y., Lau, B.P.L., Koh, Z., Yuen, C., Ng, B.K.K.: Understanding crowd behaviors in a social event by passive wifi sensing and data mining. IEEE Internet Things J. **7**, 4442–4454 (2020)

Zhu, Y., Koutsopoulos, H.N., Wilson, N.H.M.: A probabilistic Passenger-to-Train Assignment Model based on automated data. Transp. Res. Part B Methodol. **104**, 522–542 (2017). https://doi.org/10.1016/j.trb.2017.04.012

Zhu, X., Qu, W., Qiu, T., Zhao, L., Atiquzzaman, M., Wu, D.O.: Indoor intelligent fingerprint-based localization: principles, approaches and challenges. IEEE Commun. Surv. Tutor. **22**, 2634–2657 (2020). https://doi.org/10.1109/COMST.2020.3014304

Michelle Cheung is a transport policy and planning officer and is currently with the Ministry of Transport in Singapore. She received her Master of Science degree in Transport with Business Management from Imperial College London and University College London in 2020, and her Bachelor of Environmental Engineering degree from National University of Singapore in 2013. She is currently working on transport policies pertaining to future transportation modes, and holds a decade of experience with the Land Transport Authority working on rail and bus modes.

Yan Cheng is a Senior Researcher at the College of Transportation, Tongji University, China. She worked as an Associate Lecturer (Teaching) at the Centre for Transport Studies, University College London, from 2018 to 2021. Her research interests include railway planning and design, travel demand forecasting and travel behaviour analysis.

Taku Fujiyama is an Associate Professor at Centre for Transport Studies, University College London. He was trained as a railway civil engineer with much experience in station development projects. His research interest includes railway and public transport planning, operation, and management.

Authors and Affiliations

Michelle Cheung[3] · Yan Cheng[1,2] · Taku Fujiyama[3]

✉ Yan Cheng
yan_cheng@tongji.edu.cn

Michelle Cheung
michelle.cheung.19@ucl.ac.uk

Taku Fujiyama
taku.fujiyama@ucl.ac.uk

1 Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China

2 Shanghai Key Laboratory of Rail Infrastructure Durability and System Safety, Tongji University, Shanghai 201804, China

3 Department of Civil Engineering, University College London, London WC1E 6BT, UK