MIT Open Access Articles

Capacity-Constrained Network Performance Model for Urban Rail Systems

Massachusetts Institute of Technology

Capacity-Constrained Network Performance Model for Urban Rail Systems

Baichuan Mo

Research Assistant

Department of Civil and Environmental Engineering

Massachusetts Institute of Technology

77 Massachusetts Ave, Cambridge, MA 20139

Tel: +1 857-999-5906; Email: baichuan@mit.edu

https://orcid.org/0000-0003-0323-0066

Zhenliang Ma, Ph.D. (Corresponding Author)

Assistant Professor

Public Transport Research Group (PTRG)

Department of Civil Engineering

Monash University

Clayton, VIC 3800, Australia

Tel: +61 175109768; Email: mike.ma@monash.edu

Haris N. Koutsopoulos, Ph.D.

Professor

Department of Civil and Environmental Engineering

Northeastern University

Boston, MA 02115, United States

Tel: +1 617-373-6263; Email: h.koutsopoulos@northeastern.edu

Jinhua Zhao, Ph.D.

Associate Professor

Department of Urban Studies and Planning

Massachusetts Institute of Technology

77 Massachusetts Ave, Cambridge, MA 20139

Tel: +1 617-253-7594; Email: jinhua@mit.edu

Word count:5,065 words text + 3 tables × 250 words (each) = 5,815 words

Submission Date: March 1, 2020

ABSTRACT

This paper proposes a general Network Performance Model (NPM) for urban rail systems performance monitoring using smart card data. NPM is a schedule-based network loading model with strict capacity constraints and boarding priorities. It distributes passengers over the network given origin-destination (OD) demand, operations, route choice, and effective train capacity. A Bayesian simulation-based optimization method for calibrating the effective train capacity is introduced, which explicitly recognizes that capacity may be different at different stations depending on congestion levels. Case studies with data from the Mass Transit Railway (MTR) network in Hong Kong are used to validate the model and illustrate its applicability. NPM is validated using left behind survey data and exit passenger flow extracted from smart card data. The use of NPM for performance monitoring is demonstrated by analyzing the spatial-temporal crowding patterns in the system and evaluating dispatching strategies.

Keywords: Urban railway systems; effective train capacity; simulation-based optimization; network loading; performance monitoring; dispatching strategy

INTRODUCTION

Increases in ridership are outpacing capacity in many large urban rail transit systems, including Hong Kong, London, New York, and Beijing (*1*). Crowding at stations and on trains is a concern due to its impact on safety, service quality, and operating efficiency. Monitoring network performance (e.g. waiting time on platforms, load on trains, etc.) is essential to help agencies and operators in understanding the system, informing passengers, and improving operating strategies.

Compared to traditional survey methods for service performance evaluation, data from automated fare collection (AFC) and automated vehicle location (AVL) systems provide ample opportunities for analysis in areas such as travel behavior, operations planning, and monitoring, etc. (*2, 3*). For performance monitoring, some performance indicators can be directly derived from automated data, including vehicle-kilometers, vehicle-hours, travel time reliability, OD demand, etc. (*4, 5*). However, the problem of determining vehicle load and passenger waiting times is not trivial. Recently, a number of methods have been proposed to monitor passenger waiting times, left behind at stations, and vehicle loads, using AFC and AVL data (*6*). Network loading or assignment models can also be used. The main difference between network loading and assignment models lies in their behavioral assumptions. Network loading models assume that travel choices are known, while assignment models estimate the travel choices through user equilibrium criteria (network loading is a key component of transit assignment models).

Network assignment models are mainly used for planning applications. Nuzzolo et al. (*7*) proposed a dynamic schedule-based assignment model to simulate the within-day and day-to-day learning process of passengers' route choice. Nuzzolo et al. (*8*) proposed a mesoscopic transit modeling framework named "DYBUS2" to provide real-time short-term predictions of network performance. Recently, Yao et al. (*9*) developed an agent-based simulation model for the Beijing metro system. The applicability of these models for performance monitoring is limited. For performance monitoring, the interest is on the performance of the system on a particular day, which represents just one realization of operating conditions. Hence, finding an equilibrium solution is not actually applicable. Network loading models are more appropriate for this purpose.

Network loading models provide detailed performance at different levels (station, line, train, passenger) given OD flows, path choices and operating conditions. Therefore, they are suitable for modeling the performance of a particular day. For example, Grube et al. (*10*) developed an event-based network loading model to simulate metro systems in Santiago de Chile. However, since such approaches have to be applied at the network level, they require assumptions about train capacity (as assignment models also do). Determining the "actual" train capacity is not trivial. Studies found that train capacity may vary depending on the crowding levels in trains and on platforms (*11, 12*). Ma et al. (*6*) showed that ignoring the "varied" train capacity results in biased estimates of left behind (highlighting the importance of network loading models to capture this varied capacity in their representation of the system).

The paper develops a capacity-constrained, schedule-based network performance model (NPM), introduces a flexible train capacity model (called *effective* capacity hereafter), and proposes an optimization methodology to calibrate effective train capacities using AFC data. NPM models detailed passengers' trajectories, including access and egress, queuing, transferring, boarding, alighting, and left behind. The contribution of this paper is threefold:

- Develops a data-driven metro network performance model with explicit effective train capacity constraints and boarding priorities, where the effective train capacity varies by station depending on congestion levels
- Proposes a simulation-based optimization method to calibrate the effective train capacity using AFC and AVL data
- Validates the NPM and demonstrates its applicability using data from the heavily-congested metro system in Hong Kong.

1   The remainder of the paper is organized as follows: Section 2 introduces the network performance
2 model. The route choice and effective train capacity models, as well as the calibration methodology, are
3 discussed in Section 3. Case studies are presented in Section 4 to validate the NPM performance and
4 demonstrate its functionality. Section 6 concludes the paper and discusses future research directions.

5 **NETWORK PERFORMANCE MODEL**
6 Figure 1 provides an overview of the main structure of NPM. It consists of four main components: the
7 input, the network loading engine, the output of various performance indicators, and the calibration engine.
8 The calibration engine provides the capability to calibrate model parameters using reality available data.
9 Green squares in the input section represent the parameters to be calibrated. Gray squares in the output
10 section indicate model outputs that are also directly observable from AFC data (and hence, can be used for
11 calibration and validation purposes). The calibration engine compares the output journey times and OD
12 exit flows to the observed values (ground truth) and uses the difference to adjust model parameters (e.g.
13 effective capacity).
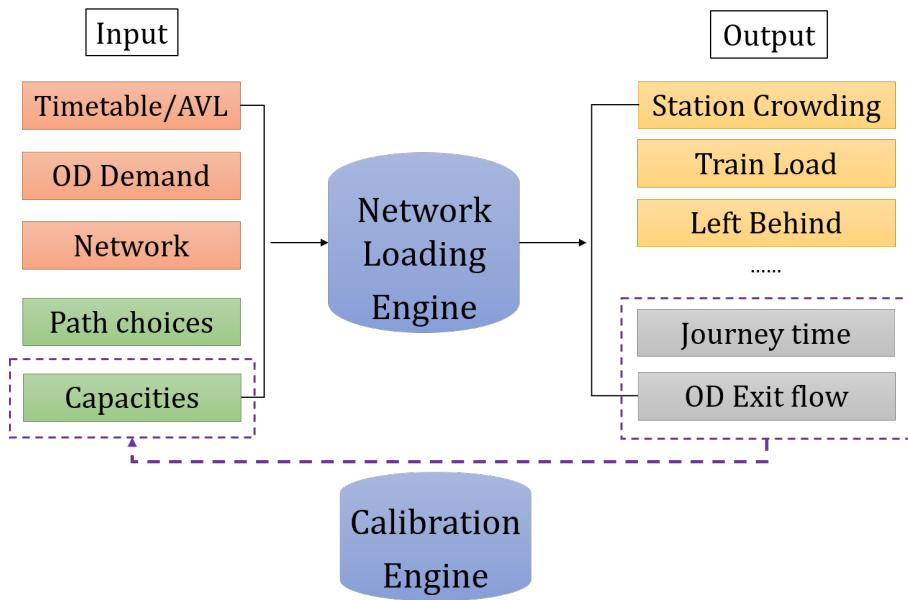
FIGURE 1: Structure of the network performance model

14 **Inputs**
15 The NPM inputs include dynamic OD demand, path choice fractions, train movement information, train
16 capacity, and access/egress/transfer walking time. Table 1 summarizes the data inputs. It is assumed
17 that the system is closed. The AFC data contains passenger's tap-in and tap-out times and stations (the
18 complete OD entry demand by time period). The timetable provides the planned train arrival and departure
19 information, while the AVL data provides the actual ones. The walk time is assumed to be normally
20 distributed with the mean and variance calculated from field observations.
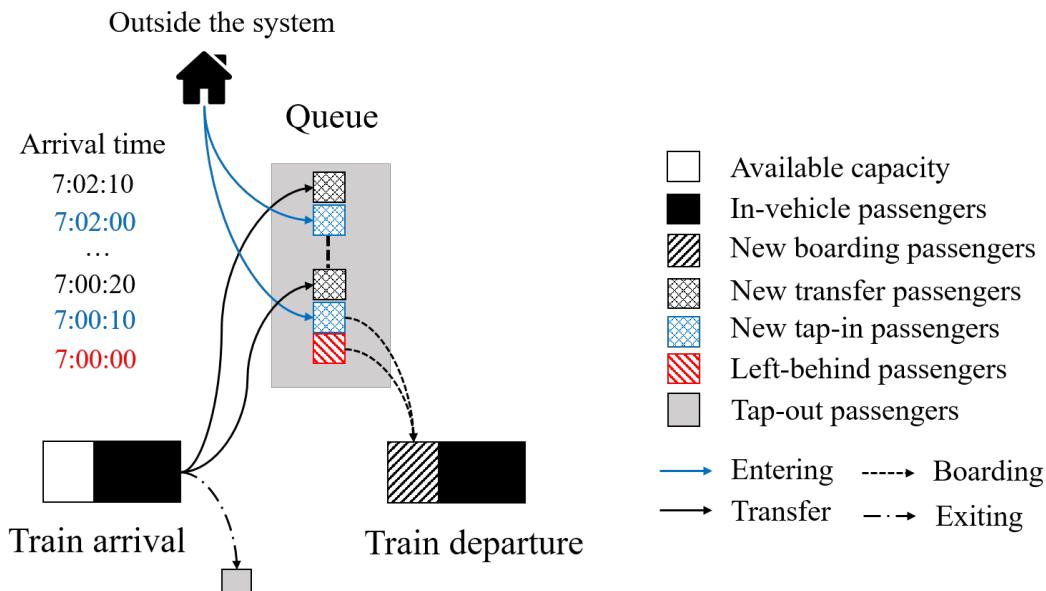
TABLE 1: Input variables and data sources

Input variables	Sources
OD entry demand	AFC data
Path choice	Section 2.2.4
Train movement	Timetable or AVL data
Train capacity	Section 2.2.5
Access/egress/transfer walk time	Measured by on-site observations or estimated from AFC/AVL data (*13*)

1 **Network loading**
2 Figure 2 summarizes the main structure of the network loading model. Three objects are defined: trains,
3 queues, and passengers. Trains are characterized by routes, runs, current locations, and capacities. Pas-
4 sengers are queued based on their arrival times. Three different types of passengers are represented:
5 left-behind passengers who were denied boarding from previous trains, new tap-in passengers from out-
6 side the system, and new transfer passengers from other lines. The left-behind passengers are usually at
7 the head of the queue.
8   An event-based modeling framework is used to load the passengers onto the network. Two types
9 of events are considered: train arrivals and train departures. The events are sorted by time and processed
10 sequentially until all events are successfully completed during the analysis period. When a train arrives
11 at a station, the offloaded passengers either transfer or exit. Transfer passengers join the boarding queue.
12 When a train departs a station, passengers are loaded on the train up to its available capacity based on a
13 First Come First Serve (FCFS) principle.

FIGURE 2: Structure of the network loading model

14 *Preprocessing*
15 Train event lists (arrivals and departures) are generated according to the actual train movement data from
16 AVL (or timetable). Each event contains a train ID, occurrence time, and location (platform). A passenger

1   is randomly assigned to a route based on the corresponding path choice probability estimated from a path
2   choice model. Random access and egress times are generated given corresponding distributions.

3   *Train arrivals*
4   For an arrival event, the train offloads passengers who reach their destination or need to transfer at the
5   station and updates its state (e.g. train load and in-vehicle passengers). For passengers who reach their
6   destinations, their tap out times are calculated by adding their egress time. For those who transfer at the
7   station, their arrival times at the next platform are calculated based on the transfer time distribution. The
8   transfer passengers are added to the waiting queue in order of their arrival times.

9   *Train departures*
10 For departure events, the queue on the platform is updated by the new tap-in passengers, that is, passengers
11 who arrive at the platform after the last train departed are added into the queue based on their arrival times.
12 Passengers board the train according to an FCFS discipline until the train reaches its capacity. Passengers
13 who cannot board are left behind and wait in the queue for the next train. The states of the train and the
14 waiting queue are updated accordingly.

15 *Route choice*
16 Route choice is usually modeled using the discrete choice framework, which assumes that decision makers
17 maximize their utilities when making choices (*14*). The multinomial logit (MNL) model is a typical
18 example of discrete choice models. For path choice problems, the C-logit model is often used. The C-
19 logit is a variation of the MNL model which corrects for the fact that alternatives may not be independent
20 due to path overlap. The C-logit incorporates an additional "cost" attribute, the commonality factor (CF),
21 in the utility (*15*). The probability of choosing path i is given by:

$$P_i = \frac{\exp\left(\beta_X \cdot X_i + \beta_{CF} \cdot CF_i\right)}{\sum_{j \in \mathbb{W}} \exp\left(\beta_X \cdot X_j + + \beta_{CF} \cdot CF_j\right)}, \tag{1}$$

22 where, X_i is the attribute vector of path i, such as in-vehicle time, number of transfers, etc. $\mathbb{W}$ is the set of
23 all alternative paths for the same OD pair. β_X and β_{CF} are the corresponding coefficients to be estimated.
24 CF_i is the commonality factor of path i, defined as:

$$CF_i = \ln \sum_{j \in \mathbb{W}} \left(\frac{L_{i,j}}{L_i L_j}\right)^{\gamma}, \tag{2}$$

25 where, $L_{i,j}$ is the number of common stations of paths i and j. L_i and L_j are the number of stations for
26 paths i and j, respectively. γ is a positive constant which is determined based on empirical studies (*16*). In
27 practice, the estimation of the route choice model parameters is based on survey data. Recent studies have
28 also proposed estimation methods based on AFC data (*17–20*).

29 *Effective capacity*
30 Train capacity is a vague concept. Normally trains may not reach their designed physical capacity for
31 various reasons (e.g. passengers may decide not to board due to the crowding (*12*)). Therefore, assuming
32 a fixed physical capacity may not be a reasonable assumption in real-world situations. In this paper,
33 we introduce the concept of *effective capacity*, which is the train capacity actually being utilized under
34 crowding situations. Effective capacity is determined by three factors: a) waiting passenger distribution
35 on the platform, b) train load and distribution across the train, and c) passengers' willingness to board a
36 crowded train. Thus, train capacity is not constant but may vary across stations. We use the term *effective*
37 *capacity* to differentiate it from the physical fixed train capacity. Effective capacity, as defined in this

1   paper, is dynamic and changes depending on the crowding state of the train and the platform.

2   Based on previous studies, two factors are included in the effective capacity (C^e) model: the current
3 train load when a train arrives at the platform (denoted as L) (*11*); and the number of queuing passengers
4 on the platform (denoted as Q) (*21*). The base capacity of train i is $C_i = \theta_0 n_i$, where n_i is the number of
5 cars of train i. C_i can be seen as the train load that represents acceptable service standards. At congested
6 stations, passengers may still board a train even if it is already crowded (*11*), which makes the actual train
7 load exceed C_i. Therefore, the effective capacity of train i at platform j ($C^e_{i,j}$) can be formulated as:

$$C^e_{i,j} = \begin{cases} \theta_0 n_i + \theta_1 L_{i,j} + \theta_2 Q_j & \text{if platform } j \text{ is in the list of congested stations} \\ \theta_0 n_i & \text{otherwise} \end{cases} \qquad \forall i, j \qquad (3)$$

8   The congested stations and time periods can be identified using AFC data (*6*). The term *platform*
9 means a combination of station+line+direction. θ_0, θ_1 and θ_2 are the parameters to be estimated. θ_0 is a
10 measurement of the service standard (passengers/car). C^e at congested stations and for congested trains is
11 expected to be higher than that of stations/trains with less crowding, hence θ_1 and θ_2 should be positive.
12 Although a linear model is used here, the proposed approach is quite general and can accommodate more
13 complex relationships between C^e and L, Q.

14 **MODEL CALIBRATION**
15 We illustrate the calibration approach using the effective capacity model, assuming two available ground-
16 truth information: observed OD exit flows and observed journey time distribution (JTD). The calibration
17 problem is formulated as an optimization problem. The objective function has two parts: the square error
18 between model-derived OD exit flows and the observations, and the difference between model-derived and
19 observed JTD. The optimization problem is formulated as:

$$\min_{\theta_1, \theta_2} \quad w_1 \sum_{i,j,t} (q^{i,j_t} - \tilde{q}^{i,j_t})^2 + w_2 \sum_{i,j,t} D_{\text{KL}}(f_{i,j_t}(x) || \tilde{f}_{i,j_t}(x)) \qquad (4a)$$

$$\text{s.t.} \quad q^{i,j_t}, f_{i,j_t}(x) = \text{Network Loading}(\theta_0, \theta_1, \theta_2) \qquad \forall i, j, t \qquad (4b)$$

20   where, q^{i,j_t} represents the number of passengers arriving from station i and exiting at station j during
21 time interval t (i.e. OD exit flows). $\tilde{q}^{i,j_t}$ is the observed OD exit flow extracted from AFC data. w_1 and
22 w_2 are the weights to balance the scale and the importance of the two parts. $f_{i,j_t}(x)$ is the probability
23 density function of the estimated JTD of passengers who come from station i and exit at station j during
24 time interval t. $\tilde{f}_{i,j_t}(x)$ the observed JTD obtained from AFC data. Eq. 4b indicates that q^{i,j_t} and $f_{i,j_t}(x)$
25 are obtained from the network loading model with θ_0, θ_1, and θ_2 as inputs. The difference of the two
26 distributions is expressed using Kullback-Leibler (KL) divergence (D_{KL}):

$$D_{\text{KL}}(f_{i,j_t}(x) || \tilde{f}_{i,j_t}(x)) = \int_x f_{i,j_t}(x) \cdot \log \frac{f_{i,j_t}(x)}{\tilde{f}_{i,j_t}(x)} dx. \qquad (5)$$

27   Solving the problem in Eq. 4 is a black-box optimization problem because of the non-analytical
28 nature of the network loading process. In this study, a Bayesian Simulation-based Optimization (BSO)
29 method (*22*) is applied. The BSO works by constructing a posterior distribution (surrogate function) that
30 best approximates the objective function. As the number of observations grows, the posterior distribution
31 improves, and the algorithm becomes more certain of which regions in the parameter space are worth
32 exploring. Given the general optimization approach, more sophisticated effective capacity models could
33 also be explored using the proposed BSO method.

1 **APPLICATIONS**
2 NPM can be used to monitor performance at four dimensions: measuring crowding, diagnosing crowding
3 causes, evaluating dispatching strategies, and evaluating network resilience.

4 **Crowding patterns**
5 Crowding is one of the most important metrics for evaluating the level of service, safety, etc. The crowding
6 indicators, directly obtained from NPM outputs, are summarized in Table 2. All indicators are time-
7 dependent with flexible aggregated intervals. The left behind rate is the probability of not boarding on
8 the first train, which can be calculated as the number of passengers who have been left behind at least
9 once divided by the total number of boarding passengers at the platform during a specific time period.
10 The number of times left behind is the number of trains missed after the first train due to crowding on the
11 trains.
12   Other service quality indicators can also be output by NPM, such as the availability of seats, the
13 number of standing passengers, and journey time reliability.

TABLE 2: Crowding indicators

Train	Train load
Platform	Waiting time
	Number of times left behind
	Left behind rate (% of passengers left behind)
	Queue length
Link	Link flow

14 **Crowding sources**
15 Since the NPM models passengers' travel behavior at the individual level, the complete trajectories of
16 all passengers are recorded. To diagnose the formation of crowding, the NPM can trace the sources of
17 passengers on each link and passengers exiting at each station. The information about where passengers
18 come from and how they contribute to loads at critical links of the system can inform operators to develop
19 specific demand management strategies, such as promotions (*2*), peak pricing, etc.

20 **Dispatching strategies**
21 Train dispatching strategies (e.g., headway adjustment, express trains) are basic instruments transit op-
22 erators use on the supply side to deal with crowding or improve service reliability. Evaluating different
23 train dispatching strategies can provide useful insights for improving service performance. NPM can be
24 used to analyze network performance under different train dispatching strategies, such as express trains at
25 different times during the peak. Such trains may skip stops in order to provide more capacity at crowded
26 stations.

27 **Network resilience**
28 Network resilience is the ability of the system to provide and maintain an acceptable level of service in
29 the face of disruptions and other challenges to normal operations, such as incidents, large scale natural
30 disasters, special events, etc. Metro systems in large cities are facing more and more service disruptions
31 due to increasing demand and aging infrastructure. These problems cause serious safety concerns and
32 service performance deterioration. The agencies are using various strategies, from demand management
33 to infrastructure improvements, to prevent disruptions and mitigate their impacts. Approaches based on
34 passenger information are still emerging as strategies transit agencies use to deal with disruptions. NPM

1 can be used to analyze network resilience by comparing the performance indicators (e.g. waiting time)
2 given different actions that operators may take when disruptions occur, such as providing information and
3 dispatching shuttle buses.

4 **CASE STUDY**
5 The NPM was demonstrated and validated using data from the MTR system in Hong Kong (Figure 3). The
6 system serves the urbanized areas of Hong Kong Island, Kowloon, and the New Territories, and consists
7 of 11 lines with 218.2 km (135.6 mi) of rail, 159 stations, including 91 heavy rail stations and 68 light rail
8 stops. The network serves over 5 million trips on an average weekday. For the urban heavy rail lines, trip
9 transactions are recorded when entering and exiting the system. The Admiralty (ADM) station is one of
10 the most crowded stations, with high volumes of passengers boarding and transferring there. In this case
11 study, the airport express and light rail transit services were not considered since they are separated from
12 the urban heavy rail lines. Passengers who enter the urban heavy rail lines from the airport express and
13 light rail services need to tap-in again.
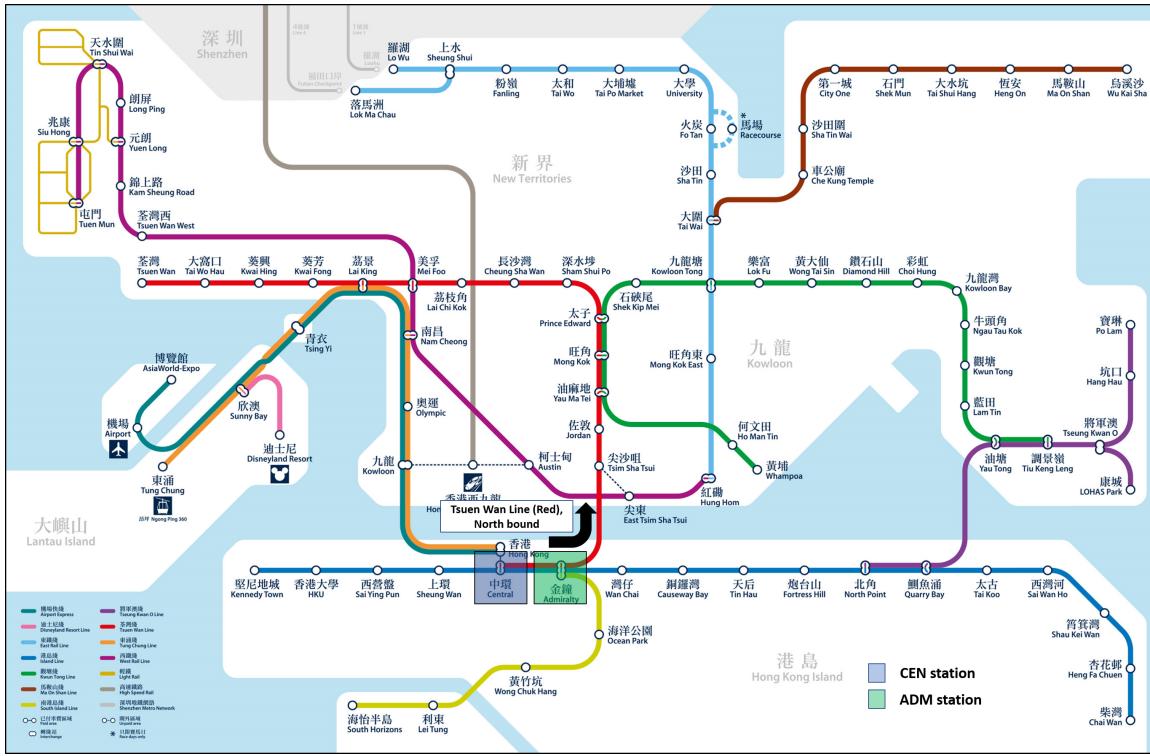
FIGURE 3: Hong Kong MTR metro system map

14 **System settings**
15 AFC data from a weekday in March 2017 is used to generate the OD entry demand and conduct effective
16 capacity calibration. Since AVL data is not available for all lines, the timetable was used to provide
17 train movement information (the actual train movements may differ from the timetable). Considering
18 the high on-time performance of the MTR system (99.9 percent on-time rate)(*23*), this is a reasonable
19 approximation. Since the evening peak is the most congested period, we only consider the period from
20 17:00 to 20:00 for model application. The warm-up and cool-down times are both set as 1 hour. The
21 running time is about 15 minutes on a personal computer with a 3.6GHz CPU and 32GB of RAM.

1 **Model Calibration**
2 The route choice model used to calculate path choice fractions for various OD pairs was estimated using
3 data from a survey of MTR users (*16*). A total of 31,640 passengers participated in the survey, with 26,996
4 valid responses. The model estimation results are shown in Table 3. The main explanatory variables are
5 the total in-vehicle time, relative walk time, and the number of transfers. The relative walk time is defined
6 as the total walk time (access + transfers + egress) divided by the map distance of the path. All variables
7 are statistically significant with expected signs. Routes with high in-vehicle, walk, and transfer times are
8 less likely to be chosen by passengers. Based on the estimated parameters, the path shares for all paths for
9 the OD pairs in the MTR system were calculated.

TABLE 3: Route choice model estimation results

	Estimate	Std. Error	t-value	
In-vehicle time	-0.147	0.011	-13.64	***
Relative walk time	-1.271	0.278	-4.56	***
Number of transfers	-0.573	0.084	-6.18	***
CF	-3.679	1.273	-2.89	**
$\rho^2 = 0.54$				

***: p<0.01; **:p<0.05.

10   The optimization problem (Eq. 4) is used for effective capacity calibration. The weights in the
11 objective function were set to $w_1 = 1$ and $w_2 = 1000$ for the error in OD exit flows and JTD, respectively.
12 The optimal coefficients are $\theta_0^* = 231.6$, $\theta_1^* = 0.0732$, and $\theta_2^* = 0.0607$. The value of θ_0^* is close to the
13 MTR standard (230 passengers/car). The signs of θ_1^* and θ_2^* are consistent with the discussion in the
14 previous section.
15   For comparison purposes, a fixed train capacity model is used as the benchmark to compare with
16 the effective capacity model. The fixed capacity for train i at platform j ($C_{i,j}^f$) is defined as:

$$C_{i,j}^f = \theta_f n_i \qquad \forall i,j \tag{6}$$

17   We test three different values of θ_f for comparison, that is, $\theta_f = 230$, $\theta_f = 245$, and $\theta_f = 260$
18 passengers/car.

19 **Results**
20 *Model validation*
21 To validate the performance of NPM, field observations at ADM station (Tsuen Wan Line, north direction)
22 on the same day as the AFC data were used for comparison. The data were collected by MTR employees
23 who counted passengers on the platform during 18:00-19:00. Left behind passengers, the number of
24 arriving passengers (sum of the new tap-in and transfer passengers) and the number of passengers boarding
25 each train were recorded.
26   Figure 4 compares the fixed-capacity model, effective capacity model, and ground truth for differ-
27 ent indicators. The number of boarding and arriving passengers from the effective capacity model matches
28 the ground truth observations well as shown in Figure 4a and 4b. The peak in Figure 4a is due to an empty
29 train dispatched from the upstream terminal station, so that more capacity is available to serve the passen-
30 gers at the crowded ADM station. The Root Mean Square Error (RMSE) of the number of boarding and
31 arriving passengers for each train is reported. The arrival passenger curves (Figure 4b) for fixed capacity
32 and effective capacity models are nearly the same. This is expected because the number of arriving pas-

sengers mainly depends on the OD demand and path shares and these two inputs are the same for the fixed and effective capacity models. However, the estimates of boarding passengers from the effective capacity model are closer to the observed values compared to the fixed capacity model (Figure 4b). The results support the importance of using the effective capacity since boarding passengers are directly related to the train capacity.

A comparison of the train load between the fixed capacity and effective capacity models is shown in Figure 4c (ground truth train load data were not available). The trains at ADM station are always full from 18:15 to 19:00. Figure 4c shows that the effective capacity can capture the variability of the train load due to the change in crowding levels over time. Other studies, based on actual observations of train loads, also support this finding (*11*).

Figure 4d compares the percentage of passengers who are left behind different times according to the models and ground-truth observations. The effective capacity model provides a more accurate estimation of left behind than the fixed-capacity model, which is consistent with findings in Ma et al. (*6*).

(a) Boarding passengers

(b) Arriving passengers

(c) Train load

(d) Left behind
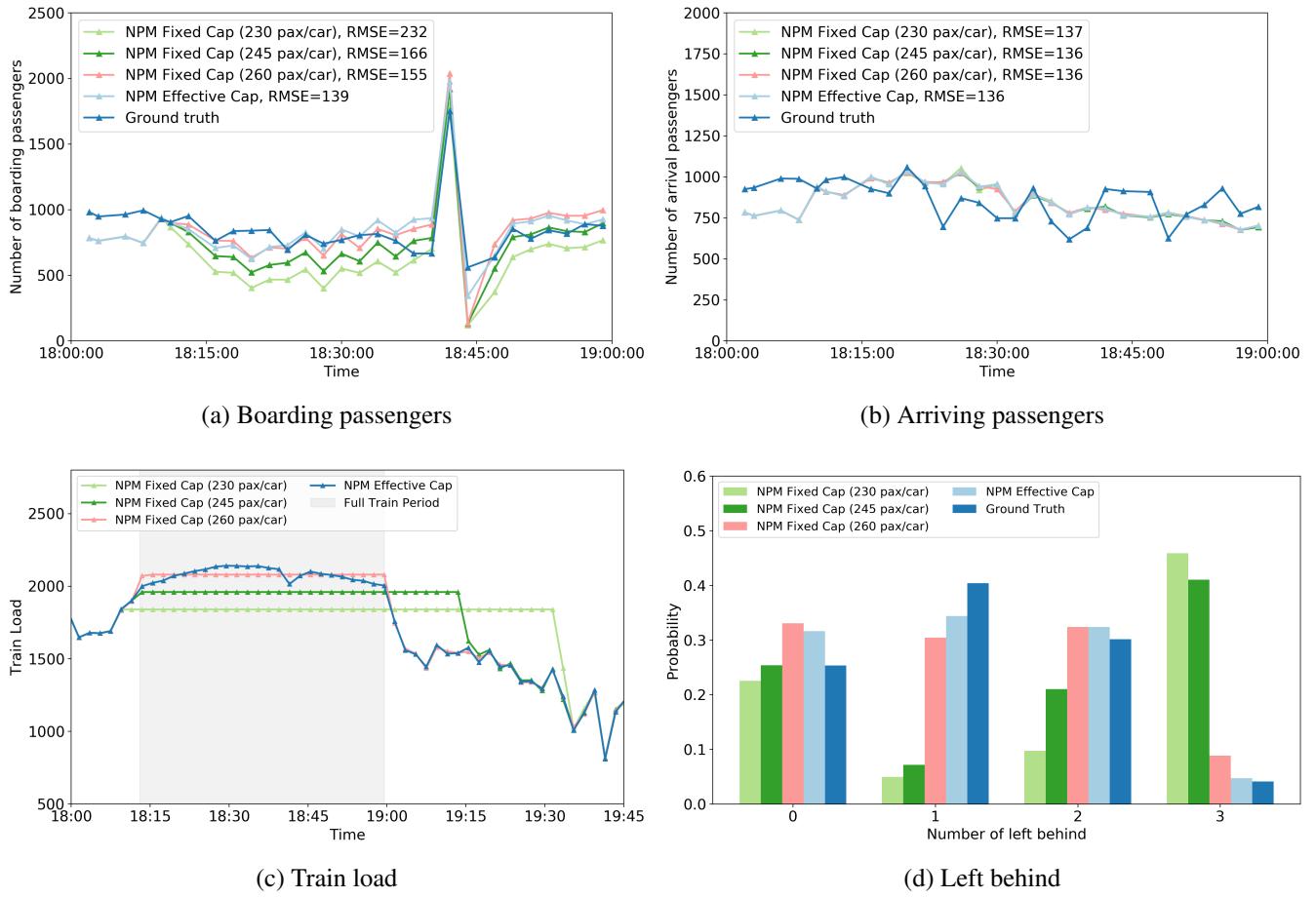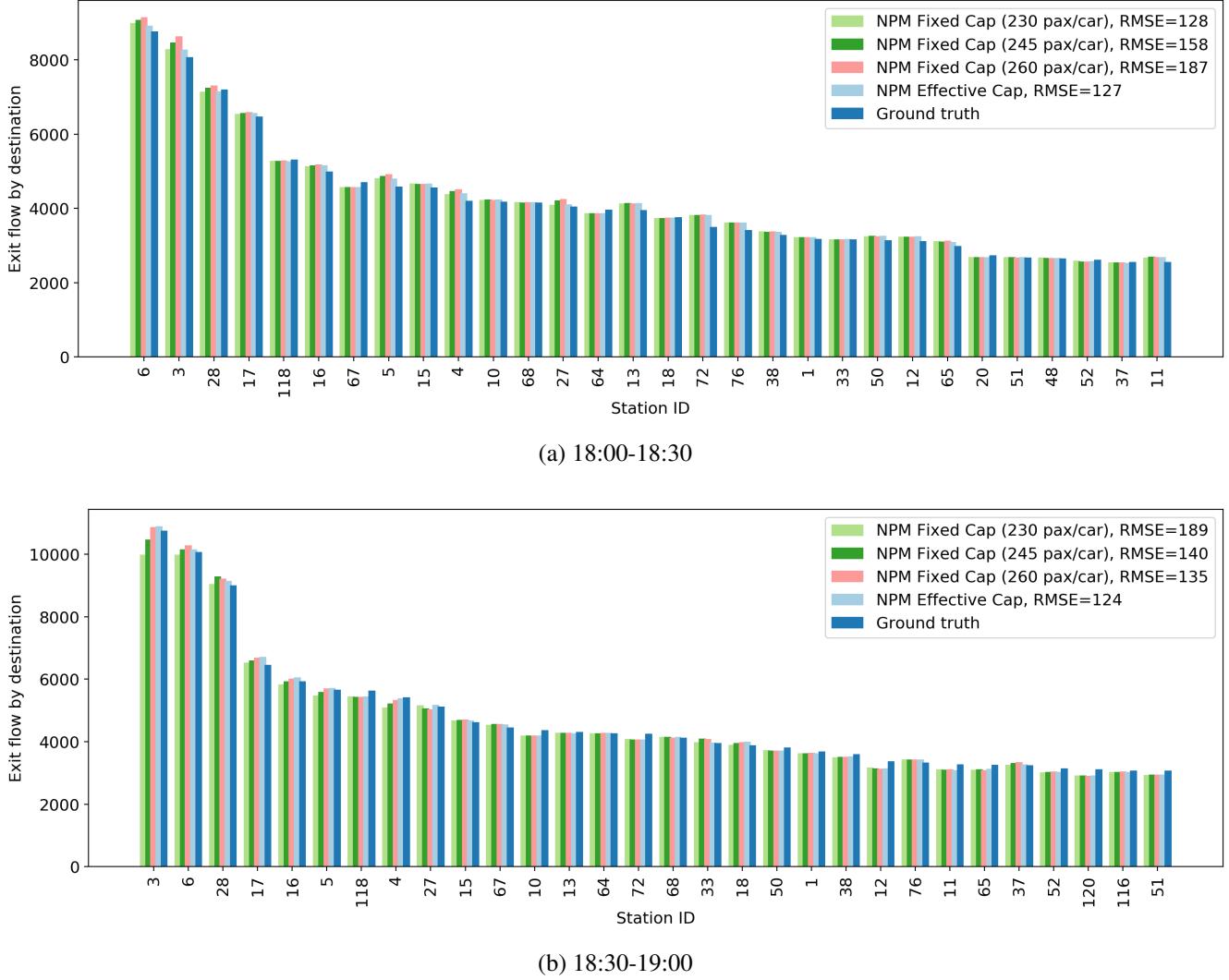
FIGURE 4: Model validation at the ADM station, Tsuen Wan Line, northbound (18:00-19:00)
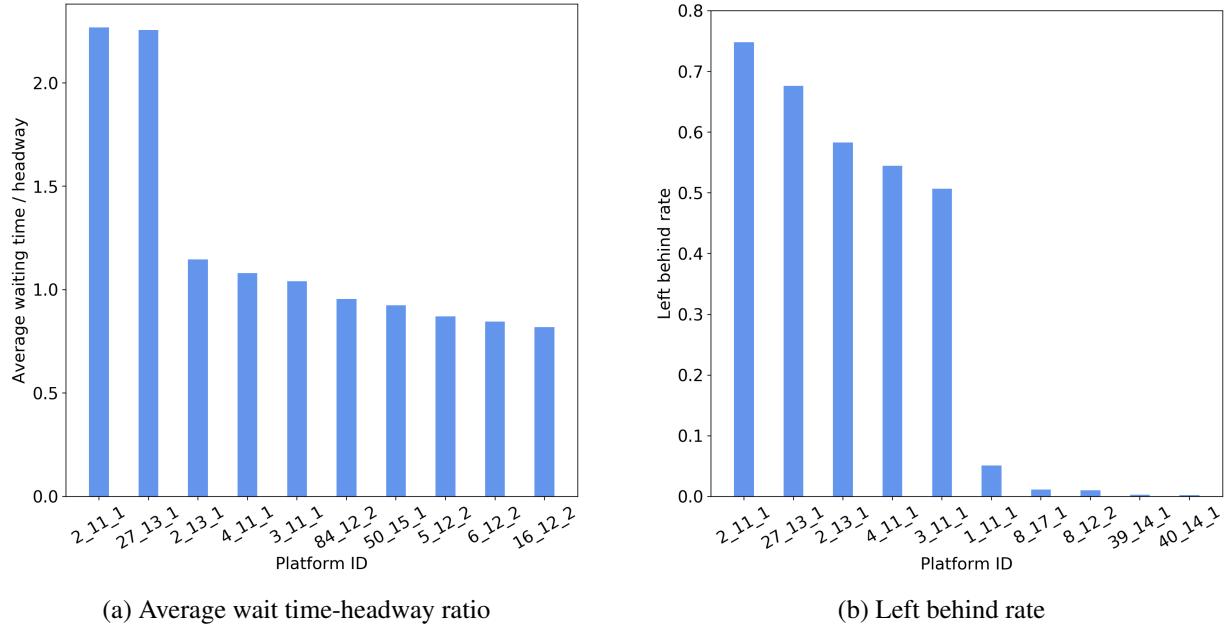
Figure 5 compares the exit flows from the NPM against the actual observations extracted from the AFC data. The top 30 stations in terms of exit flows are displayed. The RMSE of the exit flows for each model is also reported. The results from the effective capacity model match the ground truth well and outperform the fixed-capacity models.

Overall, the proposed effective capacity NPM can capture real-world situations well and has the potential to be an effective tool for performance monitoring.

(a) 18:00-18:30

(b) 18:30-19:00

FIGURE 5: Exit flow comparison (18:00-19:00)
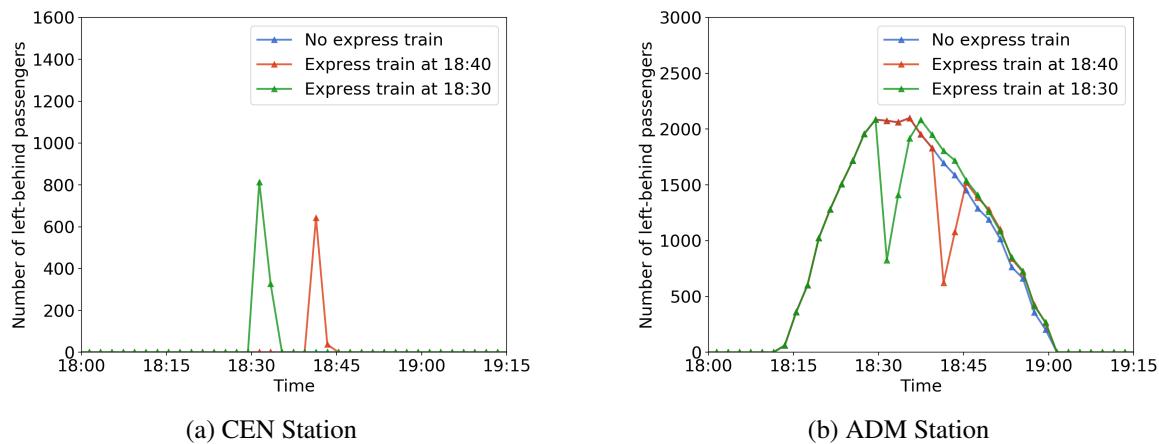
1 *Crowding analysis*
2 In a congested rail system, waiting times increase because of passengers left behind due to full trains.
3 Figure 6 shows the wait time-headway ratio and left behind rates for the top 10 crowding platforms in
4 the network. The wait time-headway ratio is defined as the passenger's average wait time divided by
5 the headway. Under normal conditions for operations with small headway variabilities and no capacity
6 constraints (assuming random passenger arrivals), the ratio has a value close to 0.5. The platform ID
7 in Figure 6 reflects the station ID + line ID + direction. For example, 2_11_1 is the platform at ADM
8 station serving the Tsuen Wan Line in the north direction. Figure 6a shows that, at platforms 27_13_1 and
9 2_11_1, passengers have to wait for an average of 2 headways. That means a passenger is expected to wait
10 for more than 2 trains to board. Figure 6b shows the top 10 platforms by their left behind rates (probability
11 to be left behind at least once). The most congested platform during the evening peak is at ADM station
12 (Tsuen Wan Line northbound). The platform has a left behind rate of about 0.75, consistent with the high
13 wait time-headway ratio shown in Figure 6a.

(a) Average wait time-headway ratio

(b) Left behind rate

FIGURE 6: Network crowding patterns (18:00-19:00)

1 *Evaluation of dispatching strategy*
2 A key application of NPM is the evaluation of different dispatching strategies. As shown in Figure 4a, an
3 empty train is dispatched (express train) from CEN to ADM at 18:40 to serve a large number of passen-
4 gers typically waiting at the ADM station. NPM can be used to test how effective such strategies are in
5 relieving congestion. For comparison purposes, we also test two additional scenarios: a) no express train
6 is dispatched; and b) the express train is dispatched at 18:30. Figure 7 compares the number of left-behind
7 passengers at CEN and ADM, which are the first two stations on the Tsuen Wan Line, northbound. Dis-
8 patching an express train transfers the congestion from ADM to CEN. The strategy temporally decreases
9 the left-behind passengers at ADM. The dispatching time does not significantly influence the crowding
10 patterns at ADM. However, the 18:30 dispatching seems to reduce the number of left behind more than
11 the 18:40 dispatching as it targets better the peak of the crowding conditions.


(a) CEN Station

(b) ADM Station

FIGURE 7: Comparison of left-behind passengers for different dispatching strategies

CONCLUSION

The paper proposed a general Network Performance Model (NPM) for network performance monitoring. The major component of NPM is an event-driven network loading module, which is capable of simulating passengers' walking, queuing, boarding, and alighting processes. NPM can be used to infer crowding patterns and evaluate dispatching strategies. An important contribution of the paper is a method for calibrating train capacity, which explicitly recognizes that capacity may be different at different stations, depending on the crowding levels on the platform and the train. The model is applied using a case study with data from Hong Kong's MTR network. The results show that NPM is able to replicate actual conditions (based on AFC data and direct observations of crowding levels at one station). NPM was also used to evaluate the effectiveness of various dispatching strategies in reducing onboard crowding. The results highlight the importance of calibrating the train capacity and support the value of the model for performance monitoring and operating strategies evaluation. Future research should focus on jointly calibrating the parameters of the path choice and train capacity models. The effective train capacity model can also be improved to better reflect passengers' willingness to board. Further sensitivity tests on the impacts of inputs accuracy could also be interesting, particularly train operations.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: B. Mo, Z. Ma, H.N. Koutsopoulos, J. Zhao; data collection: B. Mo, Z. Ma; analysis and interpretation of results: B. Mo, Z. Ma, H.N. Koutsopoulos; draft manuscript preparation: B. Mo, Z. Ma, H.N. Koutsopoulos. All authors reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

REFERENCES

1. Koutsopoulos, H. N., Z. Ma, P. Noursalehi, and Y. Zhu, Chapter 10 - Transit Data Analytics for Planning, Monitoring, Control, and Information. In *Mobility Patterns, Big Data and Transport Analytics* (C. Antoniou, L. Dimitriou, and F. Pereira, eds.), Elsevier, 2019, pp. 229 – 261.
2. Ma, Z. and H. N. Koutsopoulos, Optimal design of promotion based demand management strategies in urban rail systems. *Transportation Research Part C: Emerging Technologies*, Vol. 109, 2019, pp. 155 – 173.
3. Nassir, N., M. Hickman, and Z.-L. Ma, A strategy-based recursive path choice model for public transit smart card data. *Transportation Research Part B: Methodological*, Vol. 126, 2019, pp. 528 – 548.
4. Trépanier, M., C. Morency, and B. Agard, Calculation of transit performance measures using smartcard data. *Journal of Public Transportation*, Vol. 12, No. 1, 2009, p. 5.
5. Ma, X. and Y. Wang, Development of a data-driven platform for transit performance measures using smart card and GPS data. *Journal of Transportation Engineering*, Vol. 140, No. 12, 2014, p. 04014063.
6. Ma, Z., H. N. Koutsopoulos, Y. Chen, and N. H. Wilson, Estimation of denied boarding in urban rail systems: alternative formulations and comparative analysis. *Transportation Research Record*, 2019, p. 0361198119857034.
7. Nuzzolo, A., F. Russo, and U. Crisalli, A doubly dynamic schedule-based assignment model for transit networks. *Transportation Science*, Vol. 35, No. 3, 2001, pp. 268–285.
8. Nuzzolo, A., U. Crisalli, L. Rosati, and A. Comi, DYBUS2: a real-time mesoscopic transit mod-

eling framework. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, IEEE, 2015, pp. 303–308.

9. Yao, X., B. Han, D. Yu, and H. Ren, Simulation-based dynamic passenger flow assignment modelling for a schedule-based transit network. *Discrete Dynamics in Nature and Society*, Vol. 2017, 2017.

10. Grube, P., F. Núñez, and A. Cipriano, An event-driven simulator for multi-line metro systems and its application to Santiago de Chile metropolitan rail network. *Simulation Modelling Practice and Theory*, Vol. 19, No. 1, 2011, pp. 393–405.

11. Liu, Z., S. Wang, W. Chen, and Y. Zheng, Willingness to board: a novel concept for modeling queuing up passengers. *Transportation Research Part B: Methodological*, Vol. 90, 2016, pp. 70–82.

12. Preston, J., J. Pritchard, and B. Waterson, Train overcrowding: investigation of the provision of better information to mitigate the issues. *Transportation research record*, Vol. 2649, No. 1, 2017, pp. 1–8.

13. Zhu, Y., H. N. Koutsopoulos, and N. H. Wilson, A probabilistic Passenger-to-Train Assignment Model based on automated data. *Transportation Research Part B: Methodological*, Vol. 104, 2017, pp. 522–542.

14. Ben-Akiva, M. E., S. R. Lerman, and S. R. Lerman, *Discrete choice analysis: theory and application to travel demand*, Vol. 9. MIT press, 1985.

15. Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta, A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. In *Transportation and Traffic Theory. Proceedings of The 13th International Symposium On Transportation And Traffic Theory, Lyon, France, 24-26 July 1996*, 1996.

16. Li, W., Route and Transfer Station Choice Modeling in the MTR System, 2014, working paper.

17. Sun, L., Y. Lu, J. G. Jin, D.-H. Lee, and K. W. Axhausen, An integrated Bayesian approach for passenger flow assignment in metro networks. *Transportation Research Part C: Emerging Technologies*, Vol. 52, 2015, pp. 116–131.

18. Zhu, Y., H. N. Koutsopoulos, and N. H. Wilson, Passenger Itinerary Inference Model for Congested Urban Rail Networks. *Transportation Research C: Emerging Technologies*, 2020, under review.

19. Mo, B., Z. Ma, H. Koutsopoulos, and J. Zhao, Assignment-based Path Choice Estimation for Metro System Using Smart Card Data. In *24th International Symposium on Transportation & Traffic Theory (ISTTT)*, 2020.

20. Mo, B., Z. Ma, H. Koutsopoulos, and J. Zhao, Calibrating route choice for urban rail system: A comparative analysis using simulation-based optimization Methods. In *Transportation Rsearch Board 99th Annual Meeting*, 2020.

21. Massachusetts Bay Transportation Authority, *At What Level Does Crowding Become Unacceptable*. https://www.mbtabackontrack.com/blog/48-at-what-level-does-crowding-become-unacceptable, 2016, accessed: 2019-06-27.

22. Snoek, J., H. Larochelle, and R. P. Adams, Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, 2012, pp. 2951–2959.

23. Mass Transit Railway, Hong Kong, *MTR Maintains 99.9 percent On-Time Performance in 2017*. https://www.mtr.com.hk/archive/corporate/en/press_release/PR-18-009-E.pdf, 2018, accessed: 2020-02-18.