CrowdAtlas: Estimating Crowd Distribution within the Urban Rail Transit System

Jinlong E[†*], Mo Li[†*], Jianqiang Huang[‡*]
[†]Nanyang Technological University, Singapore [‡]Alibaba Group, China
[*]NTU-Alibaba Joint Research Institute, Singapore
ejinlong89@gmail.com, limo@ntu.edu.sg, jianqiang.hjq@alibaba-inc.com

***Abstract*—While the urban rail transit systems are playing an increasingly important role in meeting the transportation demands of people, the precise awareness of how the human crowd is distributed within the urban rail transit system is highly necessary, which serves to a range of important applications including emergency response, transit recommendation, commercial valuation, *etc*. Most urban rail transit systems are closed systems where once entered the travelers are free to move around all stations that are connected into the system and are difficult to track. In this paper, we attempt to estimate the crowd distribution within the urban rail transit system based only on the entrance and exit records of all the rail riders. Specifically, we study Singapore MRT (Mass Rapid Transit) as a vehicle and leverage the tap-in and tap-out records of the EZ-Link transit cards to estimate the crowd distribution. Guided by a key observation that the passenger inflows and arrival flows at various MRT stations are spatio-temporally correlated due to behavioral consistence of MRT riders, we design and implement a machine learning based solution, CrowdAtlas, that accurately estimates the crowd distribution within the MRT system. Our trace-driven performance evaluation demonstrates the effectiveness of CrowdAtlas.**

***Index Terms*—Urban rail transit system, Crowd distribution, Machine learning**

I. Introduction

An urban rail transit system is generally an electric railway system operating on an exclusive right-of-way [1], where passengers can travel freely among stations in different lines. In virtue of fast velocity and large capacity, the rail transit systems have become the most important urban public transportation service in recent years. Especially in many metropolises, the average daily ridership has reached millions (*e.g.,* ∼5 million in London [2] and ∼3.5 million in Singapore [3]).

Since travelers are free to move around all the stations once they enter the transit system, it is essential to have a fine estimation of how people are distributed within the transit system. Such information is important to providing critical resilience guarantees, *e.g.,* emergency evacuation in response to railway disruptions or terrorist attacks. It is also useful to other value-add businesses, *e.g.,* real-time transit recommendations based on crowdedness, or commercial valuation with crowd flows.

In this paper, we attempt to accurately estimate the crowd distribution within the urban rail transit system and take Singapore MRT (Mass Rapid Transit) system as a vehicle to study the problem. We collect one-year *EZ-Link* card data of Singapore MRT, involving daily rides (from both tap in and tap out) of 5 lines and 102 stations, totaling 1.2 billion records. With those records, we specifically study the two major lines (namely East-West line, or EW line, and North-South line, or NS line) that span across the country and possess the heaviest ridership (∼1.5 million rides take place on those two lines everyday). Fig. 1 depicts a map of the two MRT lines and the 52 affiliated stations.

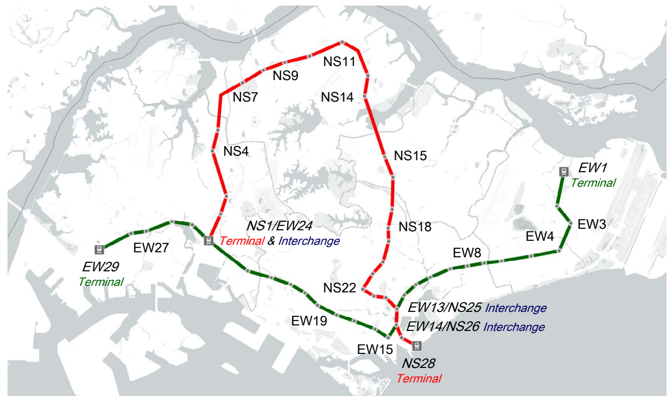


Fig. 1. A map of the two major MRT lines EW and NS.

Achieving accurate estimation of crowd distribution in the MRT, however, is challenging owing to the limited information. With the EZ-Link card data we have the knowledge of the riders that enter the MRT system, but the travelers are free to move around all the stations unspotted. Fine reasoning the crowd movement needs to overcome the uncertainties that arise from the ride time and the trip destinations. Through analysis over historical MRT trip data, we have a key observation that the passenger inflows and arrival flows at various MRT stations are spatio-temporally correlated due to behavioral consistence of MRT riders.

Guided by the observation, we design and implement a machine learning based solution, CrowdAtlas, that is able to capture MRT riders' transition probabilities and based on that perform accurate online estimation of the crowd distribution. In particular, CrowdAtlas builds a neural network model that learns the flow correlation, *i.e.,* the riders' transition probabilities among stations and across time from a high volume of historical MRT trips. With the model and all the tap-in records, the number of MRT riders at any MRT station can thus be estimated by aggregating the riders transitioned

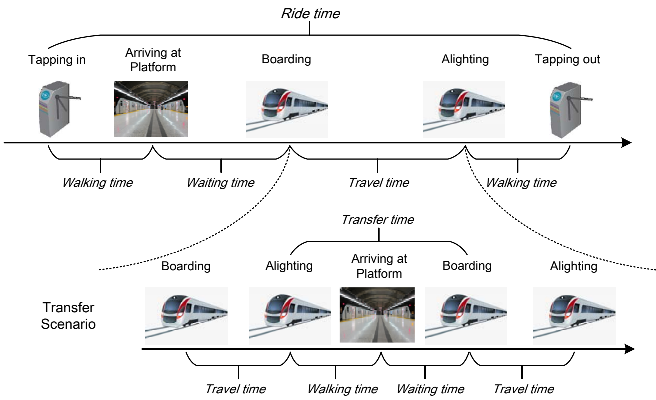
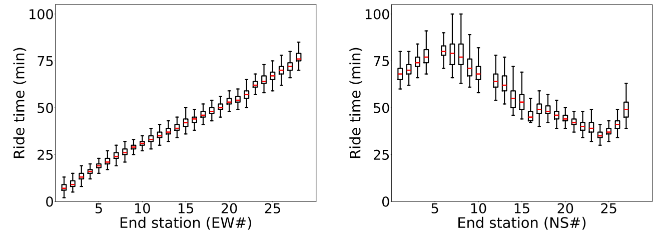


Fig. 2. Typical activities during a ride.



(a) To EW-line stations (b) To NS-line stations

Fig. 3. Ride time statistics from an MRT station EW1 to other stations.

from all origin stations and from all past time beings. We perform comprehensive evaluations with EZ-Link data traces. Our trace-driven evaluation results – the overall MAPE (Mean Absolute Percentage Error) is less than 15%, suggest that CrowdAtlas is able to produce accurate estimation of crowd distribution for most stations.

II. Problem Formulation

A passenger's rail transit ride begins from tapping in at the origin station. There are various travel and sojourn time involved during the passenger's stay within the MRT system, which finally ends when the passenger taps out at the destination station. Additionally, if the origin and destination stations are in different lines, extra sojourn time is incurred at the interchange stations. A sequence of these ride activities is depicted in Fig. 2.

Suppose there are n stations $S = \{s_j\}(j = 1, 2, \cdots, n)$ for the rail transit system, and a passenger's ride start time is τ. Based on that, a station's *inflow* is defined as the number of passengers tapping in the station at time τ, which is a τ-dependent variable. We regard the inflow set of all stations $I = \{I_j^\tau\}(j = 1, 2, \cdots, n)$ to be known, as generally these inflows could be obtained by the MRT operator in real time. Similarly, we can define a station's *outflow* $O_j^{\tau'}$ as the number of passengers tapping out the station at time τ', which is also obtained by the MRT operator every minute.

A station's *arrival flow* is defined as the number of passengers presently arriving at that station (from other stations). The present number of passengers at each station can be derived by summing inflow and arrival flow at the present time as well as the sojourn passengers' number at that station, which can be derived from previous arrival flows minus outflows (see §III-B for details). The goal is thus to estimate a set of all stations' arrival flows at present time t, *i.e.,* $A = \{A_k^t\}(t > \tau; k = 1, 2, \cdots, n)$.

Special challenge comes from the nature of our problem – to instantly estimate the crowd distribution at present time. To estimate the arrival flow at time t, we do not have passengers' tap-out records that take place after time t, *i.e.,* we only have the start (time and station) of those open trips. As a station's arrival flow is brought by other stations' inflows, we define *transition probability* $p_{oa}^{\tau t}$ which describes the probability that a passenger arrives in station s_a at time t given the ride starts from station s_o at time $\tau(< t)$. The relation between the two flow sets A and I is

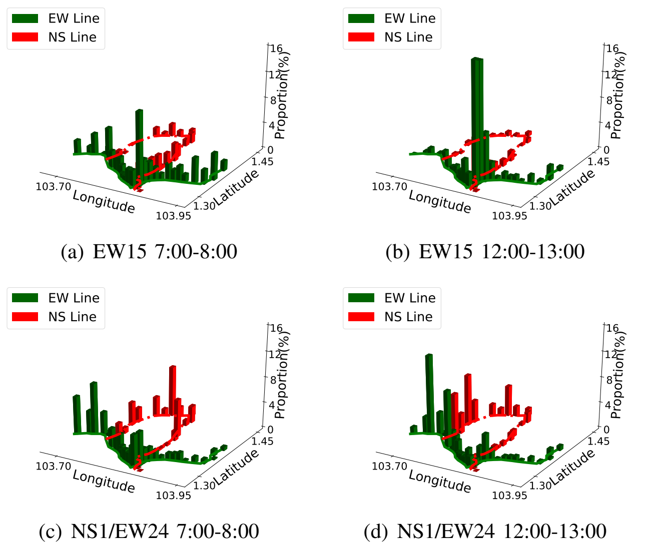
$$A_k^t = \sum_{j=1}^{n} \int_0^{t-} I_j^\tau p_{jk}^{\tau t} \, d\tau, \ \forall s_k \in S. \tag{1}$$

The transition probability for any individual passenger is affected by his/her destination station s_d as well as *ride time* T_r (*i.e.,* time spent in the rail transit system), which can be described as a function $p_{oa}^{\tau t} = f(s_d, T_r)$. It is challenging to accurately obtain $p_{oa}^{\tau t}$ due to the uncertainties of both destination s_d and ride time T_r.

We collect a large-scale EZ-Link transit card data of all Singapore MRT trips in a whole year, involving ∼2.8 million average daily ridership among 5 lines and 149 stations. We specifically study the two the major MRT lines – EW and NS lines that involve 52 stations spanning across the country with ∼1.5 million rides everyday. Here we extract one month data to illustrate and quantify the two uncertainties. Fig. 1 depicts the two MRT lines, where each station is named in the format of line name + number (*e.g.,* EW15, or NS14). Transfers may take place at 3 interchange stations marked in the figure (*i.e.,* NS1/EW24, EW13/NS25, and EW14/NS26).

Ride-Time Uncertainty. A passenger's ride time T_r comprises several time periods between ride activities – walking time, waiting time, travel time (as illustrated in Fig. 2). They are affected by either passenger behaviors (*e.g.,* walking speed, shopping activities, waiting) or train scheduling (different speeds and schedules). As a result, the ride time may vary subject to an unknown distribution ($T_r \sim F_1^u(t)$). For the above two-line dataset, the maximum, minimum and median ride times from a terminal station EW1 to all other stations are depicted in Fig. 3 respectively (separately by EW/NS line). The ride time variance is observed at each destination station with a growing trend as the station interval increases. There exists even higher variance when transfer is involved in the trip due to the additional uncertainty of transfer time.

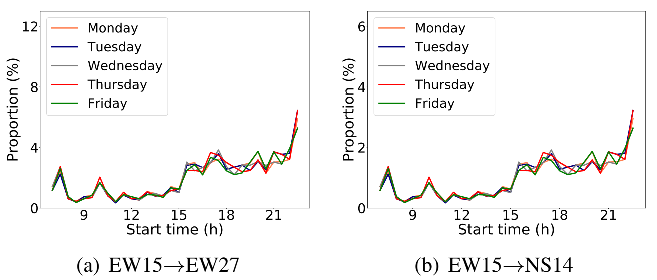
Destination Uncertainty. It is manifest that a passenger's destination station s_d is unknown until his/her ride ends. Here we concern the destinations of a batch of passengers who depart from the same origin station, and we inspect if their



(a) EW15 7:00-8:00 (b) EW15 12:00-13:00

(c) NS1/EW24 7:00-8:00 (d) NS1/EW24 12:00-13:00

Fig. 4. Destination distributions from two representative stations during peak and off-peak hours.



(a) EW15→EW27 (b) EW15→NS14

Fig. 5. Passenger transitions between (a) a pair of stations in the same MRT line, and (b) a pair of stations across different MRT lines.

distribution is analytical by studying the MRT trips. Fig. 4 depicts the destination distributions of passengers departing from two major stations of high inflows (EW15 and NS1/EW24) in a peak hour 7:00-8:00 and an off-peak hour 12:00-13:00 respectively. The proportions of passengers ending at various stations (on z axis) are projected to the map of the two MRT lines (on the x-y plane). Comparing the destination distribution from different origin stations and different start times, we clearly see that the destinations are subject to another unknown distribution among stations ($s_d \sim F_2^u(s)$), which suggests both temporal and spatial variations. In addition, interchange stations (*e.g.,* NS1/EW24) with more travel diversities further increase such uncertainty.

Key Observation. Instead of looking at the individual ride behaviors with such high uncertainties, we attempt to study the collective behaviors based on a large volume of passengers. Through quantitative analysis on a number of origin-destination station pairs, we observe spatio-temporal correlations between their inflows and arrival flows. Fig. 5 depicts the proportions of passengers transitioned from EW15 to EW27 (within the same MRT line) and to NS14 (across two MRT lines) over 5 different weekdays, where we see similar temporal variations. The observation from Fig. 5 suggests

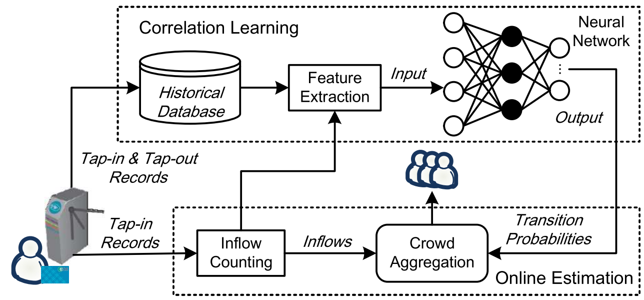


Fig. 6. Architecture of CrowdAtlas.

that the transition probability $p_{oa}^{\tau t}$ between any pair of origin-destination stations and across any time span $\tau \to t$ might be consistent across all days. In view of such an observation, we attempt to build a machine learning based model to learn such transition probabilities among stations and across time from the high volume of historical MRT trips.

III. CrowdAtlas Design

Guided by the above real-world trace analysis, we design CrowdAtlas for online estimation of the crowd distribution. Fig. 6 sketches the architecture of CrowdAtlas, which comprises two major components.

Correlation Learning: It builds a neural network model to learn the flow correlation among all stations, and consequently derive the transition probabilities $p_{oa}^{\tau t}$ between all pairs of MRT stations from inflows to arrival flows. With a list of features (start time, origin station, *etc.*) encoded as input, the model is trained based on historical MRT trip data to output transition probabilities between two stations over any time period.

Online Estimation: It takes the transition probabilities $p_{oa}^{\tau t}$ from the neural network model as input, and thus the passenger number at any MRT station (*i.e.,* crowd distribution) can be estimated in real time based on its present inflow and the arrival flow derived based on inflows from other stations and at previous time beings as well as the transition probabilities from those stations.

A. Correlation Learning with Neural Network

As mentioned above, the transition probability $p_{oa}^{\tau t}$ can be expressed as a function of two factors ride time and destination ($p_{oa}^{\tau t} = f(s_d, T_r)$), which cannot be directly derived due to the uncertainty of both factors. Instead, we design a neural network model to learn the transition probabilities between an arbitrary pair of stations and at arbitrary time of a day from flow correlation among historical MRT trip records.

Feature Extraction. $p_{oa}^{\tau t}$ is determined by both ride start and present arrival information (time and station). Here we name the above four factors as start time, origin station, present time, and present station. Note that the present station refers to the station a passenger is presently arriving at or is about to arrive at (if the passenger is presently on the MRT train). We acquire training samples by transforming the MRT trip records with ride start and end time, as well as the origin

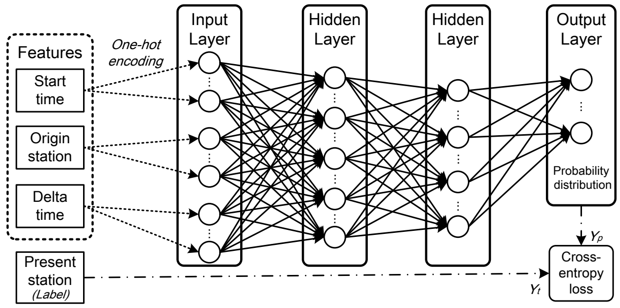


Fig. 7. The neural network model structure for transition probability learning.

and destination stations. Each MRT trip record [tap-in time, origin station, tap-out time, destination station] is transformed to a training sample of [start time, origin station, present time, present station] (where the difference between train arrival time and tapping-out time is negligible for training with high volume of data).

On this basis, we further derive *delta time* $\Delta t =$ *present time* $-$ *start time* for each record, and filter the records with $\Delta t \in [RT_{min}, RT_{max}]$ as training samples (RT_{max} and RT_{min} are the reasonable maximum and minimum ride time among stations shown in Fig. 3). For each training sample, we set (*start time*, *origin station*, *delta time*) as input features of the neural network, and take a passenger's *present station* in the record as the label. Accordingly, for a group of start time, origin station and delta time, the output of neural network is a distribution of transition probabilities to all stations that the passenger probably arrives at present. In addition, to support learning the correlation among stations in multiple lines, we re-index all stations in a sequence with all interchange stations in the front, followed by the rest of the stations by MRT lines.

Neural Network Structure. Fig. 7 depicts the neural network structure, which is composed of an input layer, an output layer and two hidden layers. To avoid the negative influence of feature values on training weight changes, we conduct *one-hot encoding* for each feature and the label. We use *ReLU* as the activation function of the hidden layers to relieve the gradient vanishing problem [4], and use *softmax* as the activation function of the output layer to generate probability distribution [5]. Meanwhile, we adopt *batch normalization* [6] after each hidden layer for faster training speed. Finally, we adopt *cross entropy* as the loss function, which indicates the closeness between two distributions [7], and set *Adam* [8] as the optimizer.

B. Online Estimation of Crowd Distribution

Through learning with the neural network, transition probabilities $p_{jk}^{\tau t}$ between any two stations $s_j, s_k \in S$ for any time τ, t during MRT operation hours can be obtained from the neural network at real time. For passengers starting from any station s_j and time τ, we can acquire the inflow I_j^τ by counting the tap-in records. Accordingly, we can estimate the distribution of passenger number $F_n(r,t)$ at an arbitrary arrival station s_r in each minute t by $F_n(r,t) = I_j^\tau p_{jr}^{\tau t}$. On this basis, we can derive the present arrival flow of a station A_r^t by aggregating the arriving passengers from different origin stations in a given duration (the concerned start time for each origin station τ_j depends on its travel time to the current station). This can be expressed as

$$A_r^t = \sum_{\tau=\tau_j}^{t-1} \sum_{j}^{n} F_n^\tau(r,t), \ \forall s_r \in S, t > \tau. \tag{2}$$

To acquire a station s_r's total passenger number, its current inflow I_r^t should be added. In addition, earlier arrived but not yet tapped out passengers are also considered. We obtain the outflow $O_r^{\tau'}$ of station s_r at time $\tau' \in [t-T_s, t-1]$ by counting the tap-out records, where T_s is the maximum sojourn time at the destination station. These tapped-out passengers are removed from the arrival flows $\{A_r^{\tau'}\}$ during the above period. Therefore, the overall crowd distribution among stations can be estimated by

$$F_N(r,t) = I_r^t + A_r^t + \sum_{\tau'=t-T_s}^{t-1} (A_r^{\tau'} - O_r^{\tau'}). \tag{3}$$

IV. Experiments

We conduct experimental study with EZ-link data traces to evaluate the estimation performance of CrowdAtlas.

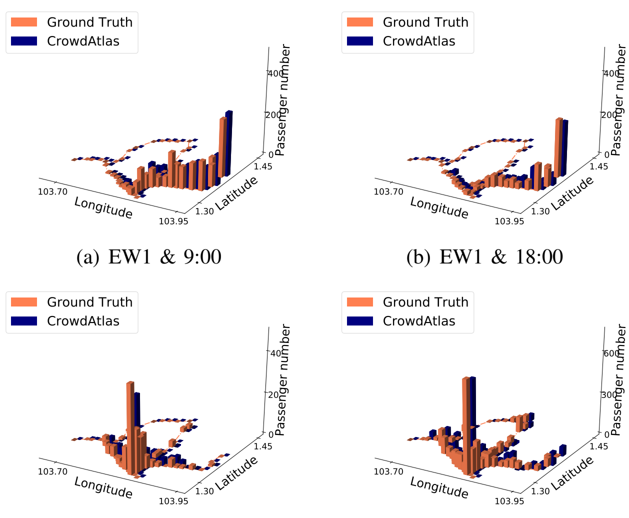
A. Data Preparation and Training

We collect one-year EZ-Link data of Singapore MRT trips, and extract a dataset of ride records within the two major lines EW and NS (as depicted in Fig. 1) for all evaluations below. To reduce the data size, we only ratain 6 major fields of each record – [origin station, start date & time, destination station, end date & time]. As the total number of stations in the two MRT lines is 52 and we conduct training in batches by the hour of start time, the unit numbers of the input and output layers are 192 and 52 respectively. We adopt an adaptive learning rate that decays as the training step number increases to avoid either fluctuation or long time learning [9]. We set the unit numbers of two hidden layers as 120 and 80, and set the typical initial learning rate and its decay rate as 0.1 and 0.95 respectively based on *grid search* [10]. Each time we take records on weekdays[1] of two months for training and records on the weekdays or weekends of the following week for testing. The selected data of different dates will be shuffled before they are input into the neural network. The training is conducted for 200 epochs so that the loss function converges.

B. Trace-Driven Performance Evaluation

We extract EZ-link data traces of the testing days and infer the passenger trajectories, through which we obtain a set of *derived* ground truths, *i.e.,* the crowd distribution among the 52 stations at any time being. The historical data provides complete start and end information of all MRT trips that take place before or after any given present time t. This allows us

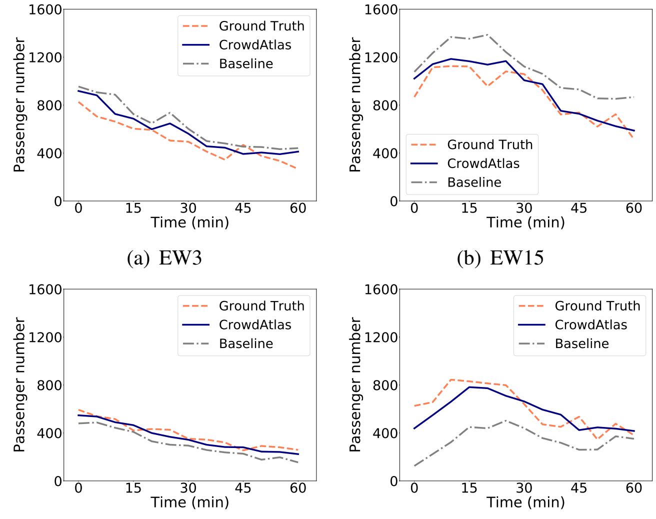
[1]The passengers' travel demand pattern on weekends might be different from that on weekdays, which could be separately trained likewise. Here we only report weekday results due to limited space.



(a) EW1 & 9:00 (b) EW1 & 18:00

(c) EW15 & 9:00 (d) EW15 & 18:00

Fig. 8. Estimated and ground-truth passenger distributions of those departing from EW1 and EW15 at 9:00 and 18:00 respectively.



(a) EW3 (b) EW15

(c) EW27 (d) NS14

Fig. 9. Passenger number variances of typical stations in different regions.

to reconstruct the ground truth of crowd distribution with no uncertainties in trip ride time and destinations.

We first study the neural network's learning performance by comparing CrowdAtlas estimation and the ground truth of passengers from a same origin station and a given start time. Fig. 8 gives the comparison results from two different origin stations EW1 and EW15 on a typical testing weekday during morning and evening peak hours of MRT. For the passengers starting from EW1 and EW15, their distributions among stations at 9:00 and 18:00 are depicted in Fig. 8(a)-8(d) respectively. The number of passengers arriving at different stations are plotted on z axis and projected to the MRT map (on the x-y plane). Both the number estimated by CrowdAtlas and that from the ground truth are depicted in each figure. We see from all figures that the estimated distributions match the ground truth very well.

We then study the temporal variation of the estimated crowd distribution. To the best of our knowledge, there are no other studies that can be applied to addressing the problem of our concern. Thus we choose to compare with an alternative approaches that we derive as a *baseline* – estimating the distribution of transition probabilities based on Moving Average (MA) [11]. Specifically, for transition probabilities $p_{jk}^{\tau t}$ of arbitrary pairs of MRT station j and k and between time τ and t, we calculate mean values of the corresponding records of the previous two-month weekdays.

We group the 52 MRT stations into 5 regions based on their geographic locations [12], [13]. In Fig. 9, we visualize the passenger number variance estimated by CrowdAtlas in comparison with that of the ground truth and baseline at four typical stations EW3, EW5, EW27, and NS14, respectively in four regions (as marked in Fig. 1). Similarly, we select the morning peak-hour period (8:00am to 9:00am) on the same testing day. We see from the four figures that the CrowdAtlas estimation curve is closest to the ground-truth

TABLE I
MAPEs (%) COMPARISON OF THE CROWD DISTRIBUTION ESTIMATION ACROSS STATIONS IN TWO TYPICAL TIME PERIODS.

Region/Line	**Stations**	**Baseline**		**CrowdAtlas**	
		30-min	**60-min**	**30-min**	**60-min**
Central Region	EW7-EW22 NS18-NS28	29.398	28.570	18.231	17.487
West Region	EW23-EW29 NS1-NS5	34.230	30.190	17.650	16.679
East Region	EW1-EW6	26.133	27.970	10.811	14.231
North Region	NS7-NS14	25.884	25.093	8.767	10.904
NE Region	NS15-NS17	20.848	21.800	9.899	11.104
EW Line	EW1-EW29	28.300	26.841	10.644	12.117
NS Line	NS1-NS28	30.659	29.310	18.572	18.080
Overall	EW1-EW29 NS1-NS28	**29.077**	**27.985**	**14.222**	**14.921**

curve, outperforming the baseline in all cases. If we define *estimation accuracy* $= 1 - |n_e - n_t|/n_t$ (n_e and n_t are the estimated and ground-truth passenger numbers respectively), the average estimation accuracy improvements over the four stations are ranging from 10.9% to 25.6%, which suggests high accuracy of CrowdAtlas estimations.

We adopt MAPE (Mean Absolute Percentage Error) to quantitatively measure the estimation accuracy in a time scale,

$$MAPE = \frac{100\%}{T} \sum_t |(y_t - \hat{y}_t)/y_t|, \quad (4)$$

where $\hat{y}_t$ and y_t ($t = 1, 2, \cdots, T$) are the estimated and ground-truth passenger numbers at different time beings. Table I gives the MAPEs statistics of both CrowdAtlas and the baseline across different station region and line groups on weekdays in 30 and 60 minutes respectively. We see that the MAPEs of CrowdAtlas are much smaller compared with the three alternative approaches for all stations, and their overall MAPEs are $\sim$14% versus $\sim$28%.

V. Related Work

There are some studies relevant to our topic, and we summarize them as follows.

Passenger Behavior Study of Rail Transit System. Previous works on the rail transit system attempted to estimate passenger's travel time [14] or plan their travel route [15] based on the transit card records. Some researchers further leverage cellular data collected from passengers' mobile phones to conduct estimation of crowd density at stations [16], [17]. All existing studies, however, are only able to extract passengers' travel demands but not their present statuses in real time. On the other hand, emergency events like railway disruptions have been studied in the transportation field, which mainly focuses on optimizing the route design of bus bridging services in presence of rail system failures [18], [19]. All existing studies assume that the crowd distributions at disrupted stations are known. Our work can produce accurate online estimation of crowd distribution within the rail transit system, which differs from existing data driven analytical studies and fills the gap of existing railway disruption studies in the transportation field.

Machine Learning for Transport Analysis. Machine learning techniques have been applied to transport analysis and prediction in recent years mainly due to their powerful capabilities in extracting hidden characteristics from historical mobility data. For instance, traffic speed prediction has been intensively studied by exploiting deep learning models including LSTM [20], CNN [21] and combination of them [22]. Likewise, traffic flow prediction has also been studied based on learning models like SAE [23] or CNN with grid partition [24]. In addition, reinforcement learning has been utilized to control traffic lights for improved road utilization [25]. Most existing studies, however, focus on road traffic learning. In contrast, our learning objective within the rail transit system is of a different purpose and more challenging, given that less knowledge on the transport operation can be extracted from the transit card data that only provides the start and end information of the trips leaving the trip details empty.

VI. Conclusion

In this paper, guided by a key observation that the passenger inflows and arrival flows at various MRT stations are spatio-temporally correlated due to the behavioral consistence of MRT riders, we design and implement CrowdAtlas, which builds a neural network model to learn the passenger transitions among stations within the urban rail transit system and based on that perform online estimation of the crowd distribution. Comprehensive performance evaluations are done with EZ-Link data traces that demonstrates CrowdAtlas's high accuracy and effectiveness.

Acknowledgment

This research is supported, in part, by NRF Singapore under its grant SDSC-2019-001, Alibaba Group through Alibaba Innovative Research (AIR) Program and Alibaba-NTU Singapore Joint Research Institute (JRI), and Singapore MOE Tier 1 grant RG18/20. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of funding agencies.

References

[1] Britannica, "Rapid Transit," 2012, https://www.britannica.com/technology/rapid-transit.

[2] TfL, "About TfL (Transport for London)—What we do," 2018, https://tfl.gov.uk/corporate/about-tfl/what-we-do.

[3] StraitsTimes, "Bus, train trips hit record high last year," 2018, https://www.straitstimes.com/singapore/transport/bus-train-trips-hit-record-high-last-year.

[4] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. of AISTATS*, 2011, pp. 315–323.

[5] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006, p. 236.

[6] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proc. of ICML*. ACM, 2015, pp. 448–456.

[7] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A Tutorial on the Cross-Entropy Method," *Annals of Operations Research*, vol. 134, no. 1, pp. 19–67, 2005.

[8] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *arXiv preprint*, 2014, p. arXiv:1412.6980.

[9] M. D. Zeiler, "ADADELTA: An Adaptive Learning Rate Method," in *arXiv preprint*, 2012, p. arXiv:1212.5701.

[10] S. M. LaValle, M. S. Branicky, and S. R. Lindemann, "On the Relationship between Classical Grid Search and Probabilistic Roadmaps," *The International Journal of Robotics Research*, vol. 23, no. 7-8, pp. 673–692, 2004.

[11] Wikipedia, "Moving Average," 2020, https://en.wikipedia.org/wiki/Moving_average.

[12] NewLaunchesReview, "Regions of Singapore," 2020, https://www.newlaunchesreview.com/regions-of-singapore/.

[13] Wikipedia, "List of Singapore MRT Stations," 2020, https://en.wikipedia.org/wiki/List_of_Singapore_MRT_stations.

[14] L. Sun, D.-H. Lee, A. Erath, and X. Huang, "Using Smart Card Data to Extract Passenger's Spatio-temporal Density and Train's Trajectory of MRT System," in *Proc. of UrbComp*. ACM, 2012, pp. 142–148.

[15] G. Wu, Y. Li, J. Bao, Y. Zheng, J. Ye, and J. Luo, "Human-Centric Urban Transit Evaluation and Planning," in *Proc. of ICDM*. IEEE, 2018, pp. 547–556.

[16] T. Holleczek, H. L. Goh, A. Spyridon, D. T. Anh, Y. Jin, S. Yin, S. Low, and A. Shi-Nash, "Traffic Measurement and Route Recommendation System for Mass Rapid Transit (MRT)," in *Proc. of SIGKDD*. ACM, 2015, pp. 1859–1868.

[17] G. Li, C.-J. Chen, W.-C. Peng, and C.-W. Yi, "Estimating Crowd Flow and Crowd Density from Cellular Data for Mass Rapid Transit," in *Proc. of UrbComp*. ACM, 2017, pp. 1–9.

[18] J. G. Jin, L. C. Tang, L. Sun, and D.-H. Lee, "Enhancing Metro Network Resilience via Localized Integration with Bus Services," *Transportation Research Part E*, vol. 63, pp. 17–30, 2014.

[19] J. G. Jin, K. M. Teo, and A. R. Odoni, "Optimizing Bus Bridging Services in Response to Disruptions of Urban Transit Rail Networks," *Transportation Science*, vol. 50, no. 3, pp. 790–804, 2015.

[20] R. Yu, Y. Li, C. Shahabi, U. Demiryurek, and Y. Liu, "Deep Learning: A Generic Approach for Extreme Condition Traffic Forecasting," in *Proc. of ICDM*. SIAM, 2017, pp. 777–785.

[21] J. Wang, Q. Gu, J. Wu, G. Liu, and Z. Xiong, "Traffic Speed Prediction and Congestion Source Exploration: A Deep Learning Method," in *Proc. of ICDM*. IEEE, 2016, pp. 499–508.

[22] Z. Lv, J. Xu, K. Zheng, H. Yin, P. Zhao, and X. Zhou, "LC-RNN: A Deep Learning Model for Traffic Speed Prediction," in *Proc. of IJCAI*, 2018, pp. 3470–3476.

[23] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic Flow Prediction with Big Data: A Deep Learning Approach," *IEEE Trasactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.

[24] J. Zhang, Y. Zheng, and D. Qi, "Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction," in *Proc. of AAAI*, 2017, pp. 1655–1661.

[25] H. Wei, G. Zheng, H. Yao, and Z. Li, "IntelliLight: a Reinforcement Learning Approach for Intelligent Traffic Light Control," in *Proc. of SIGKDD*. ACM, 2018, pp. 2496–2505.