Predicting the Impact of Disruptions to Urban Rail Transit Systems

XIAOYUN MO, CHU CAO, MO LI, and DAVID Z. W. WANG, Nanyang Technological University

2

Service disruptions of rail transit systems have become more frequent in the past decade in urban cities, due to various reasons, such as power failures, signal errors, and so on. Smart transit cards provide detailed tapping records of commuters, which enable us to infer their trajectories under both normal and disruptive circumstances. In this article, we study and predict the impact of disruptions on commuters and further evaluate the vulnerability of the rail system. Specifically, we define two metrics, stay ratio and travel delay, to quantify the impact, and we derive the predictor of each metric based on the inferred alternative route choices of commuters under disruptive circumstances. We demonstrate that the alternative route choices contribute to more similar feature distribution among different disruptions, which is crucial to tackling the main challenge of abnormal data scarcity and is beneficial for obtaining more reliable predictors for future disruptions. We evaluate our approach with a real-world transit card dataset. The result demonstrates the effectiveness of our method. Based on the predictors, we further analyse the vulnerability of the rail system. An evaluation with cross validation from taxi GPS trajectory data indicates its efficacy in discovering vulnerable rail stations as well as Origin-Destination pairs.

CCS Concepts: • **Applied computing** → **Transportation**; **Forecasting**; • **Information systems** → **Data mining;**

Additional Key Words and Phrases: Railway disruption, impact prediction, vulnerability, data scarcity

ACM Reference format:
Xiaoyun Mo, Chu Cao, Mo Li, and David Z. W. Wang. 2022. Predicting the Impact of Disruptions to Urban Rail Transit Systems. *ACM Trans. Sensor Netw.* 19, 1, Article 2 (December 2022), 17 pages.
https://doi.org/10.1145/3517015

1 INTRODUCTION

The rail system is the backbone of the **public transit systems (PTS)** in urban cities. Malfunction of the rail system even in a small region may have ripple effects and significantly impair the PTS. According to our study on Singapore **Mass Rapid Transit (MRT)** rail system, major disruptions take place due to many reasons, including technical faults, extreme weathers, human injuries, and so on. The journey of thousands or even tens of thousands of commuters may be impaired. Many of them have to abandon the PTS and resort to other transportation modes (e.g., taxis).

This work is partially supported by Singapore MOE Tier 1 Grant No. RG18/20, Singapore NRF Grant No. SDSC-2019-001, and Alibaba Group through Alibaba Innovative Research (AIR) Program and Alibaba-NTU Singapore Joint Research Institute (JRI).
Author's address: X. Mo, C. Cao, M. Li, and D. Z. W. Wang, Nanyang Technological University, 50 Nanyang Avenue, 639798 Singapore; emails: {xiaoyun001, caochu, limo, wangzhiwei}@ntu.edu.sg.

1550-4859/2022/12-ART2 $15.00
https://doi.org/10.1145/3517015

This article aims at predicting the impact of rail system disruptions at the time of occurrence. Such knowledge not only benefits the PTS provider in understanding the degradation of service and making better contingency plans with other transportation resources [22, 39] but also benefits commuters in preparing for the hazards brought by disruptions. To quantify the impact of railway disruptions on passengers, many prior studies adopt analogous measures of travel delay [13, 15, 30]. In this article, we define the following two metrics to assess the impact of disruptions on commuters. (1) **Stay ratio** indicates the percentage of rail riders during the disrupted period who choose to stay within the PTS and take alternative rail lines and/or buses to complete their trip. (2) **Travel delay** indicates the extra time spent on alternative routes for those who stay within the PTS. Higher stay ratio and lower travel delay indicate smaller impact by a disruption. With quantitative impact analysis, we further derive systematic assessment of the vulnerability of arbitrary **Origin-Destination (OD)** pairs as well as the vulnerability of rail stations. Although there have been efforts made to analysing the influence of abnormal conditions of railway on commuters [13, 30, 31], most of them apply empirical knowledge or simplified human behaviour models to reason human choices, and based on that analyze the impact on commuters. Some exploit real transportation data to understand human behaviours, but they are often limited to normal PTS conditions. In this article, adopting a unique approach, we explore the transportation data during rail system disruptions and learn from the true human choices. We train a human behaviour model from those abnormal data and apply the model to predict the impact of future disruptions.

Being simple in rationale, our approach is especially challenged by the scarcity of abnormal data, i.e., those from only six to eight major disruptions per year. A direct challenge comes from the lack of training data for us to build an accurate model using supervised learning. The limited observation of disruptions makes the trained model difficult to be reliable for future disruptions unseen in the training stage. The problem becomes more challenging if we consider that only the trips of regular commuters (which is a small portion of the total affected commuters) can be utilized to analyze human behaviours, extract features, and label impact metrics, because it is very difficult, if possible at all, to infer irregular commuters' original travel intention and thus acquire high confidence in predicting their choices under disruptions.

To address the above challenges, we intend to specially find a representative feature space where the training and test disruptions share near distributions of extracted features. Different to the situation of canonical transfer learning or feature engineering, the data in both the training and test sets in our case is scarce and hence no big picture of the distribution can be profiled. We propose a novel idea of leveraging the alternative route choices of commuters to tackle the data distribution mismatch between the training and test sets. The original training problem on the feature space relevant to disruption itself is converted to one on a different feature space relevant to alternative route choices of commuters, which unifies our view of disruptions by their effect onto commuter route choices. We demonstrate that the latter feature space refers to a much higher coverage of the full distribution by the data, and the trained model can thus be more reliable to arbitrary disruptions as long as the commuter route choices can be inferred from the disruptions. Our contributions are summarized as follows:

- To the best of our knowledge, this is the first study of impact prediction as well as vulnerability analysis under rail system disruptions that learns models from true human behaviours in disruptions.
- We propose to leverage alternative route choices of commuters to address the challenges arising from data scarcity, which enables us to build an accurate and more reliable model for arbitrary disruptions. We implement and evaluate our approach with the Singapore MRT

ride records in year 2015 that involve six major disruptions. The results demonstrate that our method outperforms all the baseline methods.

- With predicted impact factors, we assess the vulnerability of Singapore MRT rail system. A cross evaluation with the Singapore taxi data indicates the efficacy of our approach in discovering vulnerable rail stations and OD pairs.

The rest of the article is organized as follows. Section 2 introduces notations and definitions of the problem and an overview of methodology. Sections 3 to 5 detail our method for impact prediction and vulnerability analysis. We present the experimental settings and evaluation results in Section 6. Section 7 reviews the related work, after which, we conclude this article in Section 8.

2 OVERVIEW

2.1 Notations and Definitions

In this section, we introduce some notations and definitions. We first give the definitions of "disruption" and "affected OD" and then derive the two impact metrics.

We treat a rail network as a directed graph $G = (V, E)$, where V represents the set of stations and E the directed rail links between stations. A disruption can be regarded as removing some links from the graph, for an unforeseen period of time. Commuters who have travel intention across these links during the period will thus be affected. We only consider major disruptions of duration longer than 30 min for analysis, because disruptions of shorter duration are difficult for us to target affected commuters. A formal definition of disruption is given as follows.

Definition 1 (Disruption). A disruption, denoted as $e = (T, G')$, refers to a period of no train service on a set of adjacent links of G, where T is the starting time of disruption and $G' = (V, E')$ is the disrupted rail network ($E' \subseteq E$) with $E \setminus E'$ the removed links.

During the disrupted period, affected commuters either stay in the PTS (e.g., wait for service resumption, or take alternative routes from bus network and the remaining rail network) or abandon the PTS and look for other private transportation modes (e.g., taxis). To specify the influence of a disruption, first of all, we use Voronoi Diagram [34] to partition the city into Voronoi cells centering at rail stations. Each cell is a *region* containing the rail station and nearby bus stations inside the region. We can describe each of the commuter trips as an OD sample between any two of those regions. Then, we define an affected OD formally as below.

Definition 2 (Affected OD). During a disruption e, an affected OD is a pair of stations (u, v) that is unreachable in G', or is too tortuous, i.e., $d'(u, v) - d(u, v) > \lambda$, where $d(u, v)$ and $d'(u, v)$ are the number of passing stations from u to v in G and G', respectively.

We empirically set λ as 10. Affected commuters of a disruption are those with their OD being one of the affected ODs of that disruption. In this article, we aim at training two impact predictors separately for two impact metrics, namely, stay ratio and travel delay. Formally, given an affected OD (u, v) of a disruption e, the stay ratio $I_s^e(u, v)$ and travel delay $I_t^e(u, v)$ are defined as

$$I_s^e(u, v) = r_{uv}^e / r_{uv}, \tag{1}$$

$$I_t^e(u, v) = t_{uv}^e - t_{uv}, \tag{2}$$

where r_{uv}, t_{uv} are the normal ridership and travel time on the original (rail) route, and r_{uv}^e, t_{uv}^e are the disruption-affected ridership and (averaged) travel time on alternative routes in PTS, between (u, v) during the disrupted period of the day.

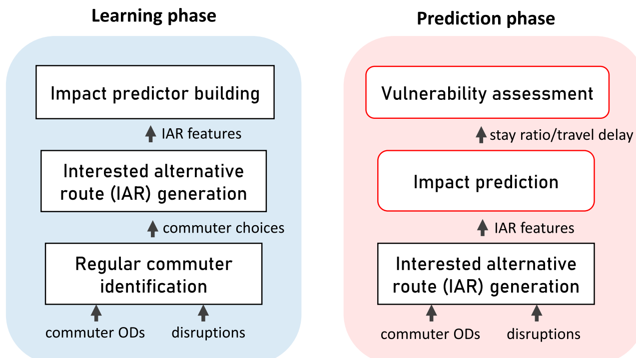

Fig. 1. Illustration of the learning and prediction phases.

2.2 Methodology Overview

We herein present an overview of our method. First, for the learning data, we were granted access to over one-year transit card records of bus and rail rides in Singapore from June 2015 to June 2016. Each record contains user ID, boarding and alighting stations and timestamps, as well as the bus/rail service name. We also obtained the information of historical disruptions from Singapore MRT operators from their official Twitter announcements that contains disruption date, starting time, ending time, and location (i.e., rail line and the stretch of disrupted stations). We specifically study the data on working days when people's travel behaviours can be stabler.

We propose the idea of utilising representative features from alternative route choices of commuters to construct impact predictors. We claim that the route choices are not only relevant to the impact of disruptions but also share similar feature distributions across disruptions, which is crucial to train reliable models in the situation of data scarcity. Figure 1 depicts the main steps of our learning and prediction phases. In the learning phase, based on historical trips, we first identify regular commuters whose travel patterns under normal condition are stable and their choices under disruption are utilized to label impact metrics and to construct impact predictors. We then generate ***interested alternative routes*** **(IAR)**, i.e., alternative routes in PTS may be chosen by commuters during a disruption, and build impact predictors based on selected features from IARs using machine learning techniques. In the prediction phase, for a given affected OD in a given disruption, the predictors can give anticipated stay ratio and travel delay as a result of impact. The two parameters are further used to analyze the vulnerability of rail stations and OD pairs.

3 REPRESENTATIVE FEATURE DOMAIN FOR SCARCE DATA

In this section, we introduce the idea of utilising alternative route choices to tackle the challenges of data scarcity. We first describe the feature domain of disruption and affected OD, compared with which, we further illustrate the proposed feature domain of alternative route choices.

Impact metrics of stay ratio and travel delay are related to features of disruption and affected OD. For example, a disruption with more broken links may lead to a lower stay ratio due to the greater mismatch in a sudden between transit demand and capacity, and a larger travel delay due to longer travelling time on alternative routes in PTS. A straightforward solution is to train a model via supervised learning based on disruption and OD features such as starting time, number of broken links, number of rail stations between an OD, and so on. Such a method, however, may result in under-fitting, since the model being trained on scarce training disruptions may not capture the functional relation, between features and impact metrics, that can extend to future disruptions.

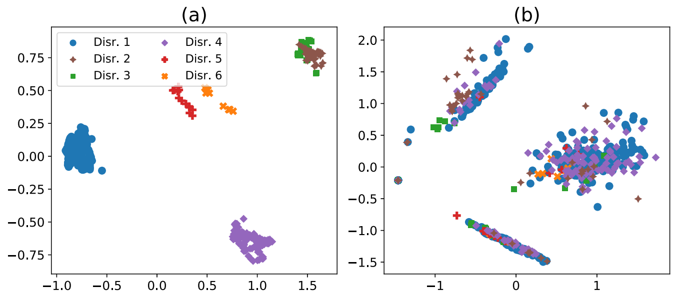

Fig. 2. Visualizing affected ODs of six observed disruptions in (a) domain $\mathbb{D}_1$ (features of disruption and OD) and (b) domain $\mathbb{D}_2$ (features of IAR).

To illustrate such a view, we name the domain of disruption and OD features as $\mathbb{D}_1 = (\mathcal{X}_1, P_1)$, where $\mathcal{X}_1$ is a d_1-dimensional feature space, and an affected OD (u, v) can be represented by a list of features $X = [x_1, \ldots, x_{d_1}] \in \mathcal{X}_1$, with a probability denoted by $P_1(X)$. We visualize all affected ODs of observed disruptions in feature space $\mathcal{X}_1$ (features are selected via the backward elimination method [17] and are listed in Table 2), after using **Principal Component Analysis (PCA)** algorithm to project them into planar points. From the result plotted in Figure 2(a), we see that distributions of points of different disruptions (colors) hardly coincide, indicating the inconsistency of the distributions in the feature space $\mathcal{X}_1$ among different disruptions. This creates a problem that a model being trained on a couple of disruptions may hardly be reliable to other disruptions. Essentially, such a problem results from the limited number of disruptions that we can observe—the data distribution of what we can observe and the data distribution of what we want to predict mismatch.

We propose to translate the prediction problem in domain $\mathbb{D}_1$ to a new domain $\mathbb{D}_2$, where the features have the following property: all observed disruptions should contain points of similar distribution. This guarantees the model applicability to future disruptions potentially with similar distribution. In our case, we aim at describing an disruption around the features of IAR, which is the alternative routes in PTS likely to be chosen by affected commuters. For an affected OD, the features of available IARs influence the choices of affected commuters, because they concern whether the IARs satisfy their requirements, e.g., a short waiting time, no transfer, and so on. The travel time of an IAR is also closely correlated to the travel delay. Therefore, we denote the new domain as $\mathbb{D}_2 = (\mathcal{X}_2, P_2)$, where $\mathcal{X}_2$ is a d_2-dimensional feature space of IAR features (listed in Table 3, which are also selected via the backward elimination method). An affected OD (u, v) can be represented by a list of features $X = [x_1, \ldots, x_{d_2}] \in \mathcal{X}_2$ with a probability denoted by $P_2(X)$. For comparison, we visualize affected ODs of observed disruptions in the feature space $\mathcal{X}_2$ in the same way as what we do in $\mathcal{X}_1$. We can see from the result shown in Figure 2(b), that the clusters of points are mixed across disruptions, suggesting highly overlapped distributions in the feature space $\mathcal{X}_2$.

Sample coverage. We further demonstrate that, given a fixed number of observations, the coverage of values in domain $\mathbb{D}_2$ by the observations is much higher than that in domain $\mathbb{D}_1$. The definition of the term *coverage* is given by Chao et al. [5]. To be specific, each feature space is divided into m equal-size blocks, and points in the same block are considered identical. Then the coverage C is defined as

$$C = \sum_{i=1}^{m} p_i \mathbb{I}[B_i > 0], \tag{3}$$

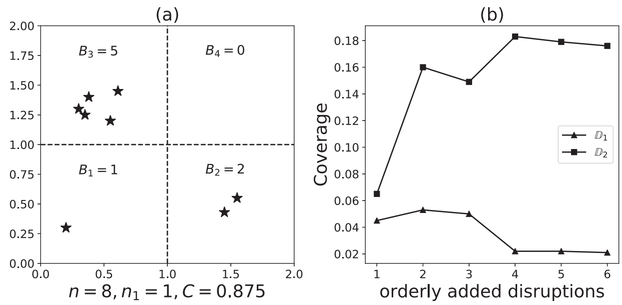

Fig. 3. (a) Illustration of the calculation of coverage C given $n = 8$, $n_1 = 1$; (b) comparison of coverages in domain $\mathbb{D}_1$ and domain $\mathbb{D}_2$ (when gradually adding new disruptions).

where B_i is the number of observations in block i, p_i is the sum of probabilities in block i (subject to the probability distribution P of the domain, e.g., $P_1(X)$ of domain $\mathbb{D}_1$), and $\mathbb{I}[\cdot]$ is the indicator function. A tractable approximation of C is given as

$$C \approx 1 - \frac{n_1}{n}, \tag{4}$$

where n is the total number of observations, and n_1 is the number of blocks, which has exactly one observation [16]. Intuitively, $1 - n_1/n$ approximates the chance that the next observation falls in a block with at least one observation before. A toy example of a 2D feature space and $m = 4$ is provided in Figure 3(a). We calculate the coverage by our data in domain $\mathbb{D}_1$ and $\mathbb{D}_2$. Feature values are normalized and each feature dimension is divided into 20 intervals. Figure 3(b) presents the relationship between coverage C and the data involving different disruptions. With six disruptions in total, we gradually add the data from each disruption. The coverage in domain $\mathbb{D}_2$ is much higher and increases much faster than that in $\mathbb{D}_1$, when involving new disruptions, which indicates that predictive models trained on $\mathbb{D}_2$ can be more reliable when applied to future new disruptions.

4 IMPACT PREDICTION

In this section, we first show the identification of regular commuters whose normal travel patterns are stable, and analyse their choices under disruptions. We then describe how we generate IAR, based on which, we build impact predictors in domain $\mathbb{D}_2$.

4.1 Obtaining Regular Commuters

Affected commuters have three kinds of choices: take alternative routes in PTS, take other transportation modes (e.g., taxis), or the mix of both. From the transit card records, we are able to find their traces in PTS during disruption. However, as travel plans of the commuters might change day by day, rendering the difficulty of deciding their original OD, it is non-trivial to infer affected commuters' choices. We propose to identify regular commuters whose travel behaviors (i.e., departure time and OD) are relatively stable and thus their original OD can be determined. After that, we focus our study on those regular commuters on behalf of all PTS users.

To identify regular commuters for a specific disruption, we obtain the list of occurrences of ODs for each commuter during the disrupted hours of the day during a sufficiently long past period (e.g., pass two months) before the date of disruption. Then the frequencies (i.e., the number of occurrences) of distinct ODs in the list are denoted by $m_1, m_2, \ldots,$ and the highest frequency is denoted by m_h. We call the OD with frequency m_h as the *dominant* OD. Then, we start to filter out commuters that are irregular. We remove those whose $m_h < max(\frac{1}{2}\sum_i m_i, \epsilon)$, to filter out occasional riders to the rail system and ϵ is empirically set to 5. We then cluster commuters with the same m_h value together. In each cluster, we remove top-25% commuters with the largest *entropy*, i.e., $-\sum_i q_i log q_i$, where $q_i = m_i / \sum_i m_i$. Note that combining both frequency (for clustering) and entropy (for intra-cluster selection) metrics, we are able to dynamically identify more regular commuters than simple thresholding approaches [23]. We assume dominant OD as the regular commuter's original travel plan, and consider only affected regular commuters whose dominant OD is one of the affected ODs.

4.2 Analysing Under-disruption Choices

Given the set of affected regular commuters, we further analyse their choices under disruptions. We find that over 90% of affected ODs have a low stay ratio, which is less than 0.5, indicating that affected commuters are prone to abandon the PTS during the disrupted period. For those who stay in PTS until arrival, over 90% of them have less than 50 min travel delay.

In addition, we analyse the alternative routes chosen by commuters who stay in PTS until arrival. An alternative route is represented by a list of bus and/or rail services, each of which is indicated by the bus service name (e.g., "Bus No.179") or a pair of rail stations. The corresponding list of transit modals for an alternative route is referred to as its *route pattern*. We find that the top-4 route patterns chosen by affected commuters, namely, (bus,), (bus, bus), (bus, rail), and (rail, bus), account for a percentage of 86%. Besides, the walking distance between two successive services is usually within 500 meters, and the detouring rate (i.e., the ratio of the distance of alternative route to that of the original route) is less than 1.5. These will serve as constraints when we generate IARs in Section 4.3.

4.3 Interested Alternative Routes Generation

For each affected OD, we first generate candidates by routing on the network of bus and disrupted rail systems using depth-first searching [27]. Candidates should satisfy constraints about walking distance (less than 500 meters between successive services), detouring rate (less than 1.5) and route pattern (being in top 7 patterns) mentioned in Section 4.2. We also gather real IARs from our transit records. We label real IARs by 1, and negative candidates that do not belong to real IARs by 0. The number of real IARs is limited (i.e., about 2.6 per affected OD) while the number of candidates is huge, leading to a significant imbalance (about 1 to 15,000) between the two classes. To alleviate the imbalance, we conduct down-sampling on the set of candidates. That is, to sample instances of candidates according to their similarity to real IARs, i.e., the more similar an instance is to any of the real IARs, the higher probability it is being sampled. To be specific, we consider several dimensions for similarity, namely, the numbers of service transfers, the number of rail stations, the number of bus stations, the length of waiting time and the distance of walking. Each dimension is normalized into the range of $[0, 1]$. Then each candidate is represented by a five-dimensional vector, and the similarity between any pair of IARs, one from real IARs and the other from candidates, is calculated by cosine similarity.

We use the features in Table 3 as well as the labels from real IARs and sampled negative candidates to train a binary classifier for IAR identification, so that for any future disruption, we can use it to distinguish real IARs from others. We train the classifier using *ensemble learning*, a supervised learning method that trains a couple of models (i.e., Decision Tree in our case for its simplicity) using different subsets of training data, and aggregates the results by majority voting. We choose ensemble learning for the purpose that it can further restrain the problem of class imbalance, as we let each subset of training data be composed of all real IARs and a different part of sampled negative candidates. For any affected OD of future disruptions, we can first generate candidate IARs and then identify real IARs using the classifier.

4.4 Predictors Building

We train the stay ratio and travel delay predictors separately. Given an affected OD, we leverage Equations (1) and (2) to calculate stay ratio and travel delay (i.e., labels). And for the input, the features of each IAR are concatenated to form a vector (inapplicable features are filled by zeros). Each IAR belongs to one group according to the length of route pattern. In each group, element-wise statistical aggregations, namely, *mean*, *max*, and *min*, over all group members are calculated, and we append the result together to form the aggregated vector. We then concatenate the aggregated vectors of all groups (in ascending order of the length of route pattern), to the form a new feature vector, where backward elimination [17] is conducted to select final vital features separately for stay ratio and travel delay. With the training samples of processed IAR features and labels, we apply **Support Vector Regression (SVR)** [12] to model the relationship between the IAR features and two impact metrics, and train the two predictors. When applying the predictors to future disruptions, we follow the same procedure of feature processing to generate testing data and predict the impact metrics for different affected ODs.

5 VULNERABILITY ANALYSIS

In this section, we provide quantitative assessment of the vulnerability of a rail station or an OD pair when it confronts a disruption, based on predicted impacts (i.e., stay ratios and travel delays).

First, we introduce some notations. We use $\mathcal{E}$ to denote the set of disruptions, and $\mathcal{E}_{uv}$ to denote the set of disruptions that affecting the OD pair (u, v). The set of ODs is represented by $\mathcal{A}$, and the set of those affected by a disruption $e \in \mathcal{E}$ is represented by $\mathcal{A}^e$. Further, we denote the set of affected ODs of disruption e originating from station u as $\mathcal{A}^e_{u*}$ ($\mathcal{A}^e_{u*} \subseteq \mathcal{A}^e$). Besides, as introduced in Section 2, the stay ratio and travel delay of affected OD (u, v) by disruption e are denoted as $I^e_s(u, v)$ and $I^e_t(u, v)$, respectively, and the normal ridership between (u, v) is denoted as r_{uv}. Then, we define the quantitative vulnerability of an OD and of a station as follows. Note that as the probabilities of occurrence of disruptions are difficult to obtain, for simplicity, we assume each disruption has equal probability of occurrence.

Vulnerability of an OD. The vulnerability of an OD pair (u, v) is defined as the averaged stay ratio (travel delay) over disruptions in $\mathcal{E}_{uv}$. Take stay ratio as an example. The vulnerability of OD of stay ratio is defined as

$$V_s(u, v) = \frac{1}{|\mathcal{E}_{uv}|} \sum_{e \in \mathcal{E}_{uv}} I^e_s(u, v). \tag{5}$$

The vulnerability of OD of travel delay $V_t(u, v)$ is defined in a similar way, replacing the $I^e_s(u, v)$ in Equation (5) by $I^e_t(u, v)$.

Vulnerability of a station. This is the one of the most concerned parameter of rail system operator.[1] Any disruption affecting a rail station influences its vulnerability. Hence, the vulnerability of a station u is defined as the averaged weighted stay ratio (travel delay) over disruptions in $\mathcal{E}$, in which the weights are set as normal riderships. For example, for stay ratio, the vulnerability of station is defined as

$$V_s(u) = \frac{1}{|\mathcal{E}|} \sum_{e \in \mathcal{E}} \left(\phi^e_s(u)\mathbb{I}[\mathcal{A}^e_{u*} \neq \emptyset] + \sigma_s \mathbb{I}[\mathcal{A}^e_{u*} = \emptyset]\right), \tag{6}$$

where $\phi^e_s(u)$ is the aggregated stay ratio over affected ODs of disruption e that are originating from station u, $\mathbb{I}[\cdot]$ is the indicator function, and σ_s is the default value if no affected OD is originating from u, which is set as 1 for stay ratio ($\sigma_s = 1$) and 0 for travel delay ($\sigma_t = 0$). Specifically, $\phi^e_s(u)$ is

[1]This is a core concern of SMRT and LTA, which are the MRT operator and transportation authority in Singapore.

Table 1. Summary of the Six Disruptions

	date	start time	duration (min)	affected line(s)	#links removed	#affected commuters	#regular commuters
1	2015-07-07	19:30	110	EW,NS	108	330,000	130,000
2	2015-10-13	08:00	30	NE	12	38,000	33,000
3	2015-10-26	05:25	90	NE	30	52,000	24,000
4	2015-11-25	05:50	140	NS	8	79,000	33,000
5	2015-12-17	19:50	135	EW	4	17,000	6,000
6	2016-03-22	11:10	160	EW	6	38,000	4,000

Table 2. Features of Disruption and Affected OD

Type	No.	Feature Description
On disruption	1	starting time, i.e., T (each time slot of 0.5 h)
	2	number of removed rail links
	3	number of affected ODs of the disruption
	4–7	binary values indicating normal/disruptive state of each of the four rail lines
On affected OD	8–11	latitude (longitude) of the origin "O" (destination "D") station
	12,13	the rail line where the origin (destination) station located
	14	number of involved stations on the original rail route between this OD
	15	number of rail stations from origin station to the nearest involved station
	16	number of rail stations from destination station to the nearest involved station

defined as

$$\phi_s^e(u) = \frac{\sum_{(u,v)\in\mathcal{A}_{u*}^e} I_s^e(u,v)\cdot r_{uv}}{\sum_{(u,v)\in\mathcal{A}_{u*}^e} r_{uv}}. \tag{7}$$

For travel delay, the vulnerability of station $V_t(u)$ is derived in a similar way, with $\phi_t^e(u)$ derived from replacing the $I_s^e(u,v)$ in Equation (7) by $I_t^e(u,v)$.

6 EVALUATION

In this section, we introduce the evaluation details of impact prediction and vulnerability analysis separately in Sections 6.1 and 6.2.

6.1 Impact Prediction

Experimental setup. We obtain the information of disruptions (e.g., time and locations) from instant tweets posted by MRT operators (i.e., SMRT and SBS).[2] We finally get six major disruptions between June 2015 to June 2016, detailed information of which are summarized in Table 1. These disruptions occurred at different locations of different rail lines ("EW" for East-West Line and "NS" for North-South Line). Considering the starting time and duration, two of them took place during peak hours,[3] and four took place during off-peak hours, and their duration vary from 30 to 160 min. In addition, to identify regular commuters and generate alternative routes, PTS static metadata (i.e., bus/rail service routes, station ids, and locations) are obtained from the LTA Online Datamall [7], and walking/riding distance between two geographical locations are acquired via Google's Direction APIs [26].

[2] Twitter accounts `SMRT_Singapore` and `SBSTransit_Ltd`.
[3] Morning peak: 07:30–09:30; Evening peak: 17:30–19:30.

Table 3. Features of Interested Alternative Route

Type	No.	Feature Description
route -wise	1	number of bus/rail services of the route
	2	normal ridership of the first service
	3	number of rail stations travelled in the original rail route
service -wise	4,5	the shortest (longest) waiting time for the service
	6	walking (access) distance
	4	a binary value indicating bus/rail transit modal of the service
	7	number of bus (or rail) stations traveled
	8,9	normal boarding (alighting) ridership of the service around starting time T near the origin station
	10	number of other bus services that can board from nearby (<500 m) bus stations

For comparison, we conduct experiments for the following methods:

- $\mathbb{D}_1$-SVR: which is built on domain $\mathbb{D}_1$ of features in Table 2 of disruption and affected OD. SVR is applied for modeling and backward elimination is used before training the model.
- PIRD: **P**redicting **I**mpact of **R**ail **D**isruptions, the method proposed in this article. IAR features are listed in Table 3.
- $\mathbb{D}_2$-Oracle: which is built on domain $\mathbb{D}_2$ and is implemented the same way as PIRD, but with feature input from real IARs (i.e., those alternative routes truly selected by commuters during disruptions, which are not available during prediction and can only be obtained after the commuters complete their trips). This approach utilizes practically not available information and its performance represents the upper-bound for comparison.
- PIRD-LR: which is implemented the same way as PIRD, except that the final regression model applied to IAR features is linear regression instead of SVR.
- $\mathbb{D}_2$-SVDD: which implements the same approach as PIRD except that it selects candidate IARs using a binary classifier named SVDD [33], an ***One-class Classification*** **(OCC)** algorithm, instead of the ensemble learning used in PIRD. As OCC uses only one class of positive samples, it avoids the imbalance issue but loses information from negative samples.

We adopt a *leave-one-out* scheme to evaluate the proposed impact predictors. Each time, we take affected OD samples from five disruptions as training set and samples from the remained disruption as test set, representing the same setting when we apply our solution in reality, i.e., we have historical disruptions to train predictors for a future disruption. The hyper-parameters are tuned using fivefold cross validation on the training set. We run the experiment for each of the evaluated methods 200 rounds. We use **Mean Absolute Error (MAE)** to evaluate the performance averaged over all six events. For each method, we calculate an *average* MAE and a *worst* MAE over all tested ODs from six disruptions. Specifically, we denote the absolute error of the jth OD in the kth round when using the ith disruption as test data as $AE_{jk}^{(i)}$, for $k = 1, \ldots, 200$, $i = 1, \ldots, 6$ and $j = 1, \ldots, J_i$, where J_i is the number of affected ODs tested in the ith disruption. Then the *average* MAE is derived as $\sum_i \sum_k \sum_{j=1}^{J_i} AE_{jk}^{(i)} / (200 \sum_i J_i)$ and the *worst* MAE is derived as $Max_i(\sum_k \sum_{j=1}^{J_i} AE_{jk}^{(i)} / (200 J_i))$.

Prediction accuracy. Both the *average* and *worst* MAE of evaluated methods on stay ratio and travel delay prediction are provided in Table 4. As we can see, PIRD outperforms $\mathbb{D}_1$-SVR. For stay ratio, PIRD provides 0.11 *average* and 0.12 *worst* MAE, while $\mathbb{D}_1$-SVR gives 0.16 *average* and 0.22 *worst* MAE. In travel delay prediction, PIRD gives 11.9 min *average* and 14.5 min *worst* MAE,

Table 4. MAE Comparison of Evaluated Methods

Impact metric	$\mathbb{D}_1$-SVR		$\mathbb{D}_2$-SVDD		PIRD-LR		PIRD		$\mathbb{D}_2$-Oracle	
	avg.	*worst*	*avg.*	*worst*	*avg.*	*worst*	*avg.*	*worst*	*avg.*	*worst*
Stay ratio	0.16	0.22	0.13	0.14	**0.11**	0.14	**0.11**	**0.12**	**0.11**	0.14
Travel delay (min)	13.8	15.5	11.8	14.9	11.4	15.1	11.9	14.5	**10.2**	**11.2**

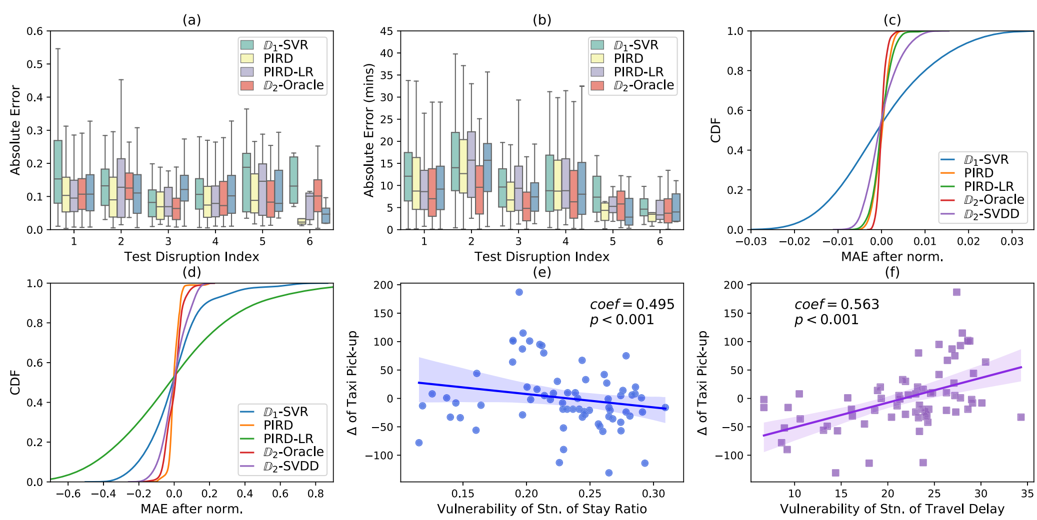

Fig. 4. (a, b) Statistics of prediction errors for different disruptions' ODs; (c, d) stability comparison between evaluated methods for stay ratio (c) and travel delay (d); (e, f) vulnerability values vs. taxi pick-up changes of all rail stations.

while $\mathbb{D}_1$-SVR gives 13.8 min *average* and 15.5 min *worst* MAE. The results suggest the training performance over domain $\mathbb{D}_2$ outperforms that conducted on $\mathbb{D}_1$. PIRD achieves MAEs close to that of $\mathbb{D}_2$-Oracle in stay ratio prediction, which indicates our IAR generation and identification methods work well in producing real choices of commuters. The results also suggest PIRD has close *average* MAEs to PIRD-LR's in both stay ratio and travel delay prediction, but outperforms PIRD-LR in the *worst* MAE, which suggests the performance gain of applying SVR over linear regression.

Figures 4(a) and (b) present statistics of prediction errors for different ODs across different disruptions. X-axes indicate test disruptions, y-axes indicate the absolute errors for (a) stay ratio and (b) travel delay. We have two observations from these two figures. First, PIRD performs close to $\mathbb{D}_2$-Oracle for most disruptions and may even outperforms it (e.g., for the Disruption 6 due to the fact that there are insufficient regular commuters and very few real IARs that can be used by $\mathbb{D}_2$-Oracle). Second, PIRD performs well in generalization. In the data records, Disruption 1 has 354 affected ODs, which account for 60% of the total number of ODs in the study. When we use Disruption 2–6 as training data, and build the model to predict the impact to ODs in Disruption 1, however, we see that the error (i.e., an average of 0.1 for stay ratio and 9 min for travel delay) is not apparently higher than what we can obtain for other disruptions. It suggests that PIRD is able to capture critical features and has strong ability in generalization with small training data.

We also compare the performance of PIRD and that of $\mathbb{D}_2$-SVDD. Specifically, for stay ratio, PIRD provides smaller *average* and *worst* MAEs (0.11 and 0.12) than $\mathbb{D}_2$-SVDD (0.13 and 0.14). For

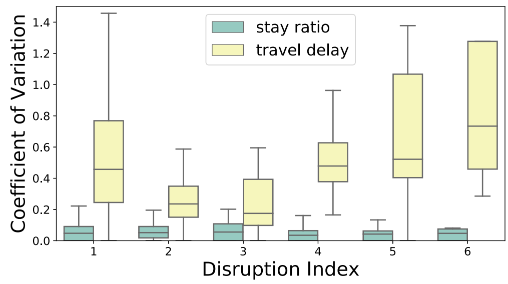

Fig. 5. Coefficient of variations of stay ratio and travel delay for affected ODs.

travel delay, PIRD achieves very close performance to $\mathbb{D}_2$-SVDD but our method leverages much less data for training, accounting for only 4.4% of all IAR candidates while $\mathbb{D}_2$-SVDD's training data usage nearly doubles our data, i.e., 8.6%. The results indicate the ensemble learning design in PIRD can effectively select meaningful IARs that are possibly chosen by commuters during a disruption.

Stability. Stability provides tolerance to perturbations from training data, which is significant to our problem as perturbations may come from data noises, variation from new disruptions, emergent actions taken during the disruptions, and so on. We evaluate the stability of PIRD in comparison with other methods. Stability can be reflected from the results of predictors. Figures 4(c) and (d) present the distributions of MAE (averaged across six disruptions) from each round of prediction after zero-mean normalization, where each line represents 200 MAE points. The sharper the ascending curve is, the more consistent the MAEs are, indicating a stabler output. Figure 4(c) shows that PIRD is stabler than $\mathbb{D}_1$-SVR and $\mathbb{D}_2$-SVDD, and is close to the performance of $\mathbb{D}_2$-Oracle. With regard to the travel delay in Figure 4(d), PIRD is stabler than other methods, and even outperforms $\mathbb{D}_2$-Oracle mainly due to the fact that $\mathbb{D}_2$-Oracle has insufficient real IARs for training stable models. The results suggest the high stability of PIRD as compared with other methods.

Limitation in obtained ground truth. We inevitably involve noises when fetching ground truths of stay ratio and travel delay, which can be introduced by the very difference of commuters' choices under one affected OD, or the fluctuation of ridership in historical days. We use ***coefficient of variation*** **(CoV)** to measure the dispersion of each OD. CoV is defined as the ratio of the standard deviation to the mean value, and is a standard measurement for dispersion. When labelling the stay ratio, the normal ridership (averaged over historical days) has an average CoV of 5.6%. For travel delay, the delays of IARs chosen by commuters for an OD usually cover a wide range, and hence the standard deviation is much higher, leading to an average CoV of 90%. Figure 5 plots the distribution of CoV for stay ratio and travel delay across all the disruptions. It clearly shows that CoV of travel delay is much higher than stay ratio for all disruptions. As a result, the averaged travel delay is less representative for all commuters in the ground truth itself, while stay ratio is a stabler and more truthful impact metric that has higher consistency across individual commuters. This is probably one of the reasons why the performance of PIRD in travel delay prediction is not as good as that of stay ratio based on the imperfect "ground truth."

6.2 Vulnerability Analysis

Experimental setup. We refer to an independent dataset that provides us knowledge of taxi pick-ups with time and locations, and use it to cross validate the vulnerability analysis results of PIRD. We were granted access to a taxi trajectory dataset from July to August, 2015. The biggest of our observed disruption, Disruption 1, took place during this time. We use the change of taxi

Table 5. The Pearson Correlation Between Taxi Pick-up Changes and Derived Vulnerability Values

	Number of objects	Stay ratio		Travel delay	
		coefficient	p-value	coefficient	p-value
Vulnerability of Station	82	−0.495	<0.001	0.563	<0.001
Vulnerability of OD	1,807	−0.256	<0.001	0.12	<0.001

pick-ups in the Voronoi region of a rail station to validate the vulnerability metrics that we obtain. The dataset were collected from more than 12,000 taxis, covering the entire Singapore. Every taxi reported its GPS location and status (vacant or busy) every 30 s to traffic authority. There are about 10 million records per day. We find that the overall number of taxi pick-ups has no significant difference between the disruptive and normal days, which is 16,200 during Disruption 1 and 16,600 (average over one month) during the same period on normal days. In small regions, however, the number of taxi pick-ups has obvious variations. The rationale is a more vulnerable station may yield higher taxi pick-ups during the disruption as more commuters choose to abandon the PTS. A less vulnerable station may yield less taxi pick-ups than usual as taxis are absorbed by those more vulnerable regions where drivers are easier to get a business. The same rationale applies to validating the vulnerability of a given OD. Within the region of rail station u, we denote the average taxi pick-ups on normal day as $vol(u)$, and the taxi pick-ups during Disruption 1 as $vol'(u)$. We use the change of taxi pick-ups $\Delta vol(u) = vol'(u) - vol(u)$ to compare with the vulnerability values of stations, i.e., $V_s(u)$ and $V_t(u)$, that PIRD derives. Similarly, we use $\Delta vol(u, v) = vol'(u, v) - vol(u, v)$ to represent the change of the number of taxi trips from the region of station u to that of station v, to compare with the vulnerability values of of ODs, i.e., $V_s(u, v)$ and $V_t(u, v)$, that PIRD derives.

Cross validation results. Table 5 quantifies the relation between the derived vulnerability values, and the change of taxi pick-ups (i.e., $\Delta vol(u)$) or the change of the number of taxi trips (i.e., $\Delta vol(u, v)$), using Pearson correlation coefficient. There are 82 stations and 1,807 OD pairs considered for calculation. For the vulnerability of station, stay ratio and the change of taxi pick-up are negative related with Pearson coefficient being −0.495, while travel delay and the change of taxi pick-up are positive related with Pearson coefficient being 0.563, both suggesting high correlations. For the vulnerability of OD, there are similar correlations with the change of the number of taxi trips but not as strong as those for the vulnerability of station, which is because the number of taxi trips matched with a specific OD is much smaller than that matched with a station, thus subject to higher deviations. In Figures 4(e) and (f), we visualize the values of vulnerability of station *v.s.* the changes of taxi pick-ups for each rail station. We see the Pearson fitting with them and the high consistency across the data. *p-value* in Table 5 indicates the significance of such correlations.

Top-20 vulnerable rail stations. We show the top-20 rail stations as suggested by the lowest vulnerabilities of stay ratio Top^{20}_{stay}, as well as those of travel delay Top^{20}_{delay}. Figure 6 depicts those top-20 stations (black circle for Top^{20}_{stay}, and black dot for Top^{20}_{delay}). We find high consistency between the two sets. Nineteen of the 20 stations are commonly agreed by the two sets. Many of those vulnerable stations are distributed outside the core downtown areas, probably due to the insufficient alternatives PTS can supply. Interestingly, some suburban stations on “NE” line in the northeast are not recognized as vulnerable stations, probably because there are two local Light Rapid Transit lines that enrich the transit supply in that area. Based on this founding of vulnerability assessment, we suggest that, in short term, the operators should put more emphasize on suburban stations when delivering emergent measures during a disruption. And in long term, the authority may invest more to improve the PTS, e.g., adding direct bus lines, in suburban area.

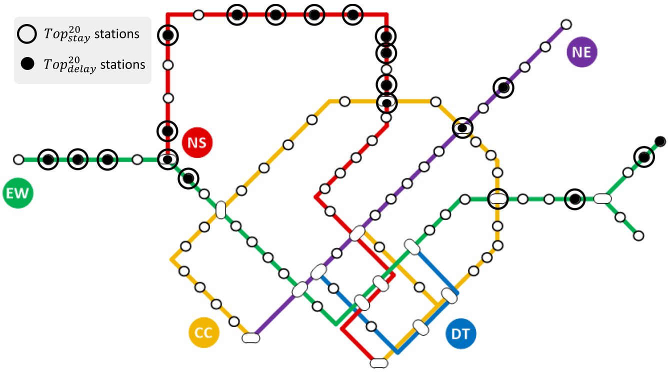

Fig. 6. Visualization of the most vulnerable rail stations.

7 RELATED WORK

In this section, we introduce priors works on impact prediction of transportation incidents, vulnerability analysis, as well as technologies relevant to our approach.

Impact Prediction. Except for studies regarding to the detection of abnormal circumstances of transportation incidents, e.g., railway failures, traffic congestion [10, 14, 28, 32], there have also been efforts made to predict their impact. Some works predict the impact by reasoning human reactions or the damage to network structure, most of which lack measurement study of real incidents. For examples, Sun et al. [31] estimates the real-time spatio-temporal distribution of commuters in rail system under normal conditions, and try to infer the number of affected commuters when there is a disruption. Sun et al. [30] try to reason commuters' travel delay based on their choices (e.g., stay or leave PTS). Yin et al. [36] define the impact as the damage to rail network efficiency, and utilize graph theory to quantify the impact of disruption. A few works make impact prediction based on actual mobility data measured from real world. Examples include Pan et al. [24] who take the average impact of similar historical incidents to predict that of future incidents, Fang et al. [13] who leverage contextual features and post-incident travel delays to predict future travel delays, and Garib et al. [15] who use statistical models based on contextual features to predict travel delay. Sometimes they require the set of contextual features to be informative, which limits the model utility, e.g., authors in References [2, 20] tried to predict how long a traffic incident would last. Most existing studies are not validated with real world incidents at the scale of this article. Other studies focus on forecasting the traffic flow under anomalous conditions [4, 11, 37] taking a period of post-incident traffic flow as input. The traffic flows, however, cannot be translated to fine-grained impact to commuters. To sum up, so far there is no existing study that measures impact from real incidents, and meanwhile explores the reliability of models being trained on scarce data to predict the impact of a variety of future incidents.

Vulnerability Analysis. Existing studies on vulnerability (or resilience) analysis intend to measure the performance of a system when suffering abnormal incidents. Some early studies such as Bruneau et al. [3] and Adams et al. [1] define resilience by introducing the *resilience triangle*, which measures resilience using severity of disruption and duration to recovery. As a rail network can be regarded as a complex network, graph theories and algorithms can also be applied to define resilience. For example, Sun et al. [29] define the vulnerability of a rail station as the connectivity reduction when it was removed from the network. Yang et al. [35] propose station importance and robustness using topological features (e.g., node degree). These studies measure vulnerability in an abstract way, which is not validated in practical conditions. Some other studies define vulnerability in a more realistic way, by associating it with the influence to commuters, e.g., alternative routes [6] and travel delay [9]. There are also some studies exploring how to improve the vulnerability of a rail system. Jin et al. [18] propose to slightly modify existing bus services to satisfy sudden demand due to accidental events. Laporte et al. [21] propose to maximize the proportion of commuters that have alternative routes inside the rail system, when designing a new rail network. None of existing studies consider actual impact prediction that is based on real-world measurement study and validated by data from true incidents.

Feature manipulation. The idea of finding a feature domain of alternative route choices to tackle challenges of data scarcity in this study shares similarity with some other feature manipulation techniques, e.g., *feature engineering* [19]. The data distribution mismatch between training and test sets however is not the focus of feature engineering. The most relevant study is that of *transductive transfer learning* [25], which transfers knowledge from the training set to the test set when the data distribution of their feature spaces are different. Existing studies such as Zadrozny et al. [38] and Daume III et al. [8] use re-sampling or statistical adaptation of the distribution of training set to that of the test set. All existing studies assume ample labeled test data, and that the global distribution of population can be profiled, which is distinct from our case.

8 CONCLUSION

We propose a comprehensive solution to predict the impact of rail system disruptions, based on the real behaviors of affected commuters during disruptions. Two metrics, stay ratio and travel delay, are defined to quantify the impact. To tackle the challenge of training data scarcity, we propose to project a disruption and its affected OD into a different domain of features abstracted from commuters' alternative route choices. The training accuracy and generalizing ability can be greatly improved. Experimental results using real-world data demonstrate the effectiveness of the proposed solution. Our further vulnerability analysis based on impact metrics identifies most vulnerable stations within the Singapore MRT rail network, which is cross-validated with taxi pick-up changes between normal and disruptive circumstances.

REFERENCES

[1] Teresa M. Adams, Kaushik R. Bekkem, and Edwin J. Toledo-Durán. 2012. Freight resilience measures. *J. Transport. Eng.* 138, 11 (2012), 1403–1409.
[2] Stephen Boyles, David Fajardo, and S. Travis Waller. 2007. A naive Bayesian classifier for incident duration prediction. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board.* Citeseer.
[3] Michel Bruneau, Stephanie E. Chang, Detlof Von Winterfeldt, et al. 2003. A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra* 19, 4 (2003), 733–752.
[4] Manoel Castro-Neto, Young-Seon Jeong, Myong-Kee Jeong, and Lee D. Han. 2009. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Syst. Appl.* 36, 3 (2009), 6164–6173.
[5] Anne Chao and Shen-Ming Lee. 1992. Estimating the number of classes via sample coverage. *J. Amer. Stat. Assoc.* 87, 417 (1992), 210–217.
[6] Lichun Chen and Elise Miller-Hooks. 2012. Resilience: An indicator of recovery capability in intermodal freight transport. *Transport. Sci.* 46, 1 (2012), 109–123.
[7] Singapore Land Transport Authority Datamall. Retrieved from https://www.mytransport.sg/content/mytransport/home.html.
[8] Hal Daume III and Daniel Marcu. 2006. Domain adaptation for statistical classifiers. *J. Artific. Intell. Res.* 26 (2006), 101–126.
[9] Alicia De-Los-Santos, Gilbert Laporte, Juan A. Mesa, and Federico Perea. 2012. Evaluating passenger robustness in a rail transit network. *Transport. Res. Part C: Emerg. Technol.* 20, 1 (2012), 34–46.
[10] Anuj Dimri, Harsimran Singh, Naveen Aggarwal, Bhaskaran Raman, K. K. Ramakrishnan, and Divya Bansal. 2019. Barosense: Using barometer for road traffic congestion detection and path estimation with crowdsourcing. *ACM Trans. Sensor Netw.* 16, 1 (2019), 1–24.

[11] Daizong Ding, Mi Zhang, Xudong Pan, Min Yang, and Xiangnan He. 2019. Modeling Extreme Events in Time Series Prediction. In *Proceedings of the 25th ACM Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD'19)*. 1114–1122.

[12] Harris Drucker, Christopher J. C. Burges, Linda Kaufman, Alex J. Smola, and Vladimir Vapnik. 1997. Support Vector Regression Machines. In *Advances in Neural Information Processing Systems*. MIT Press, 155–161.

[13] Zhihan Fang, Yu Yang, Desheng Zhang, et al. 2019. MAC: Measuring the Impacts of Anomalies on Travel Time of Multiple Transportation Systems. *Proceedings of ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 42.

[14] Fausto Pedro García Márquez, Clive Roberts, and Andrew M. Tobias. 2010. Railway point mechanisms: Condition monitoring and fault detection. *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit* 224, 1 (2010), 35–44.

[15] A. Garib, A. E. Radwan, and H. Al-Deek. 1997. Estimating magnitude and duration of incident delays. *J. Transport. Eng.* 123, 6 (1997), 459–466.

[16] Irving J. Good. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* 40, 3–4 (1953), 237–264.

[17] Isabelle Guyon and André Elisseeff. 2003. An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3 (Mar. 2003), 1157–1182.

[18] Jian Gang Jin, Loon Ching Tang, Lijun Sun, and Der-Horng Lee. 2014. Enhancing metro network resilience via localized integration with bus services. *Transport. Res. Part E: Log. Transport. Rev.* 63 (2014), 17–30.

[19] Samina Khalid, Tehmina Khalil, and Shamila Nasreen. 2014. A survey of feature selection and feature extraction techniques in machine learning. In *Proceedings of the Science and Information Conference*. IEEE, 372–378.

[20] Woon Kim, Suhasini Natarajan, and Gang-Len Chang. 2008. Empirical analysis and modeling of freeway incident duration. In *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*. IEEE, 453–457.

[21] Gilbert Laporte, Angel Marin, Juan A. Mesa, and Federico Perea. 2011. Designing robust rapid transit networks with alternative routes. *J. Adv. Transport.* 45, 1 (2011), 54–65.

[22] Zhidan Liu, Zengyang Gong, Jiangzhou Li, and Kaishun Wu. 2020. Mobility-aware dynamic taxi ridesharing. In *Proceedings of the IEEE 36th International Conference on Data Engineering (ICDE'20)*. IEEE, 961–972.

[23] Lakmal Meegahapola, Thivya Kandappu, Archan Misra, et al. 2019. BuSCOPE: Fusing individual & aggregated mobility behavior for "Live" smart city services. Retrieved from https://arXiv:1905.06116\/.

[24] Bei Pan, Ugur Demiryurek, Cyrus Shahabi, and Chetan Gupta. 2013. Forecasting spatiotemporal impact of traffic incidents on road networks. In *Proceedings of the IEEE 13th International Conf. on Data Mining*. IEEE, 587–596.

[25] Sinno Jialin Pan and Qiang Yang. 2009. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22, 10 (2009), 1345–1359.

[26] Google Maps Platform. Retrieved from https://cloud.google.com/maps-platform/.

[27] Robert Sedgewick. 2002. *Algorithms in C, Part 5: Graph Algorithms*. Addison-Wesley Professional, Boston, MA.

[28] Rijurekha Sen, Abhinav Maurya, Bhaskaran Raman, Rupesh Mehta, Ramkrishnan Kalyanaraman, and Amarjeet Singh. 2014. Road-RFSense: A practical RF sensing–based road traffic estimation system for developing regions. *ACM Trans. Sensor Netw.* 11, 1 (2014), 1–32.

[29] Daniel Sun, Yuhan Zhao, and Qing-Chang Lu. 2015. Vulnerability analysis of urban rail transit networks: A case study of Shanghai, China. *Sustainability* 7, 6 (2015), 6919–6936.

[30] Huijun Sun, Jianjun Wu, Lijuan Wu, Xiaoyong Yan, and Ziyou Gao. 2016. Estimating the influence of common disruptions on urban rail transit networks. *Transport. Res. Part A: Policy Pract.* 94 (2016), 62–75.

[31] Lijun Sun, Der-Horng Lee, Alex Erath, and Xianfeng Huang. 2012. Using smart card data to extract passenger's spatio-temporal density and train's trajectory of MRT system. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*. ACM, 142–148.

[32] Hanqing Tao and Yang Zhao. 2015. Intelligent fault prediction of railway switch based on improved least squares support vector machine. *Metal. Min. Industry* 7, 10 (2015), 69–75.

[33] David M. J. Tax and Robert P. W. Duin. 2004. Support vector data description. *Mach. Learn.* 54, 1 (2004), 45–66.

[34] Georges Voronoi. 1908. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les parallélloèdres primitifs. *Journal für Die Reine und Angewandte Mathematik* 134 (1908), 198–287.

[35] Yuhao Yang, Yongxue Liu, Minxi Zhou, Feixue Li, and Chao Sun. 2015. Robustness assessment of urban rail transit based on complex network theory: A case study of the Beijing subway. *Safety Sci.* 79 (2015), 149–162.

[36] Haodong Yin, Baoming Han, Dewei Li, et al. 2016. Evaluating disruption in rail transit network: A case study of Beijing subway. *Procedia Eng.* 137 (2016), 49–58.

[37] Rose Yu, Yaguang Li, Cyrus Shahabi, Ugur Demiryurek, and Yan Liu. 2017. Deep learning: A generic approach for extreme condition traffic forecasting. In *Proceedings of the SIAM International Conference on Data Mining*. SIAM, 777–785.

[38] Bianca Zadrozny. 2004. Learning and evaluating classifiers under sample selection bias. In *Proceedings of the 21st International Conference on Machine Learning*. 114.

[39] Pengfei Zhou, Yuanqing Zheng, and Mo Li. 2012. How long to wait? Predicting bus arrival time with mobile phone based participatory sensing. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*. 379–392.

Received 5 October 2021; revised 30 January 2022; accepted 3 February 2022