MAC: Measuring the Impacts of Anomalies on Travel Time of Multiple Transportation Systems

ZHIHAN FANG, Rutgers University, USA
YU YANG, Rutgers University, USA
SHUAI WANG, Southeast University, China
BOYANG FU, Rutgers University, USA
ZIXING SONG, Southeast University, China
FAN ZHANG, SIAT, Chinese Academy of Sciences & Shenzhen Beidou Intelligent Technology Co., Ltd., China
DESHENG ZHANG, Rutgers University, USA

Urban anomalies have a large impact on passengers' travel behavior and city infrastructures, which can cause uncertainty on travel time estimation. Understanding the impact of urban anomalies on travel time is of great value for various applications such as urban planning, human mobility studies and navigation systems. Most existing studies on travel time have been focused on the total riding time between two locations on an individual transportation modality. However, passengers often take different modes of transportation, e.g., taxis, subways, buses or private vehicles, and a significant portion of the travel time is spent in the uncertain waiting. In this paper, we study the fine-grained travel time patterns in multiple transportation systems under the impact of urban anomalies. Specifically, (i) we investigate implicit components, including waiting and riding time, in multiple transportation systems; (ii) we measure the impact of real-world anomalies on travel time components; (iii) we design a learning-based model for travel time component prediction with anomalies. Different from existing studies, we implement and evaluate our measurement framework on multiple data sources including four city-scale transportation systems, which are (i) a 14-thousand taxicab network, (ii) a 13-thousand bus network, (iii) a 10-thousand private vehicle network, and (iv) an automatic fare collection system for a public transit network (i.e., subway and bus) with 5 million smart cards.

CCS Concepts: • **Networks** → **Sensor networks**; • **Human-centered computing** → *Ubiquitous and mobile computing design and evaluation methods*; Empirical studies in ubiquitous and mobile computing;

Additional Key Words and Phrases: travel time components, anomalies, cyber physical systems

ACM Reference Format:
Zhihan Fang, Yu Yang, Shuai Wang, Boyang Fu, Zixing Song, Fan Zhang, and Desheng Zhang. 2019. MAC: Measuring the Impacts of Anomalies on Travel Time of Multiple Transportation Systems. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 2, Article 42 (June 2019), 24 pages. https://doi.org/10.1145/3328913

Authors' addresses: Zhihan Fang, Rutgers University, Piscataway, NJ, 08854, USA, zhihan.fang@cs.rutgers.edu; Yu Yang, Rutgers University, Piscataway, NJ, 08854, USA, yu.yang@rutgers.edu; Shuai Wang, Southeast University, Nanjing, Jiangsu, China; Boyang Fu, Rutgers University, Piscataway, NJ, 08854, USA; Zixing Song, Southeast University, Nanjing, Jiangsu, China; Fan Zhang, SIAT, Chinese Academy of Sciences & Shenzhen Beidou Intelligent Technology Co., Ltd. Shenzhen, Guangdong, China; Desheng Zhang, Rutgers University, Piscataway, NJ, 08854, USA, desheng.zhang@cs.rutgers.edu.

2474-9567/2019/6-ART42 $15.00
https://doi.org/10.1145/3328913

1 INTRODUCTION

According to the United Nations, we are undergoing a rapid process of urbanization where 54% of the world's population has already been moved into urban areas in 2014, and this number is projected to rise to 70% by 2050 [33]. Thus, it is essential to improve the mobility of urban residents on a daily basis, which can be achieved by accurately providing travel time estimations for improving passenger confidence when selecting different transportation systems, e.g., subway, taxi, bus, and private vehicles. However, urban anomalies, e.g., transportation accidents [38] and social events [35], have major impact on travel time across all transportation systems in cities. In this paper, our goal is to understand, quantify, and predict the impact of urban anomalies on the travel time of different transportation systems, which is essential to many real-world applications, e.g., emergency response [15], trip planning [16], and location-based services [33].

In recent years, as a result of urbanization, urban transportation systems have been equipped with advanced sensing and communication devices to generate massive amount of data [14], which provide an excellent opportunity for travel time estimation. Researchers have accumulated abundant knowledge for travel time modeling through various state-of-the-art models [22, 35, 37, 38]. In this paper, we focus on a combination of three key factors as follows to advance the state-of-the-art models. **(i) Fine-Grained Travel Time**: since most transportation data are mainly collected for management purposes and do not provide direct measurement of different stages of a trip, most existing studies have been focused on end-to-end travel time [22, 37], instead of fine-grained travel time. **(ii) Urban Anomalies**: the existing studies on urban anomalies were mainly focused on the anomaly detection from traffic flows [35, 38] instead of measuring the impact of anomalies on travel time. **(iii) Multiple Transportation Systems**: Due to limited data access, almost all existing work has been focused on an individual transportation modality, e.g., taxi [24, 26] or subway [7], instead of multiple transportation systems.

To the best of our knowledge, little work, if any, has been conducted under the above three factors. This is because it is challenging to access large-scale data on urban mobility and anomalies with fine-grained spatiotemporal coverage across different transportation systems. In the vision of smart cities, many of them, e.g., New York City, Beijing and Shenzhen, have been collecting urban-scale data across different systems to improve urban efficiency. Some of these data have been made available for researchers to understand urban mobility, e.g., travel time[37]. However, since most of these data are collected for billings across different systems, they lack direct measurement of detailed components of the travel time, e.g., walking time and waiting time.

In this paper, to address the above challenges, we design a measurement framework called MAC to Measure the impact of Anomalies on different travel time Components of heterogeneous transportation systems. In particular, we utilize various existing data sources (e.g., vehicle GPS, fare transactions, etc) to infer walking time, waiting time, and riding time for taxi, bus, subway and private vehicles. As a result, the key novelty of MAC is that it measures the impact of urban anomalies on fine-grained travel time components of multiple transportation systems, by utilizing data already collected for billing and management purposes. The key contributions of the paper are as follows:

(1) We utilize various transportation infrastructures and their data for travel time measurement under urban anomalies. To our knowledge, the utilized data is fairly complete, i.e., including taxi cab, bus, subway and private vehicle data for the same city Shenzhen. The data covers more than 78% of 11 million permanent residents. We further collect a city-scale dataset of urban anomalies including expected and unexpected anomalies as ground truth for analyses. More importantly, for the benefit of the IMWUT community, we will release our sample data.
(2) We design a framework called MAC to investigate the fine-grained travel time components and the impact of anomaly events. In particular, we design a model to infer the waiting time and riding time and then validate the inferred travel time components through case studies. Furthermore, we analyze the travel time

patterns on the inferred travel time components under normal and anomaly events. Finally, we design a learning model to integrate contextual information for the prediction of delay time in anomalies across different transportation systems.

(3) We implement and evaluate MAC by integrating data from four transportation systems including (i) a 15-thousand taxicab network, (ii) a 13-thousand bus network, (iii) an automatic fare collection system for a public transit network (i.e., subway and bus) with 5 million smartcards, and (iv) a 10-thousand private vehicle network. Our prediction model achieves 86.5% prediction accuracy on delay time prediction. Our research efforts lead to several insights and lessons learned, which are helpful for understanding city-scale fine-grained travel time under extreme anomaly condition across different transportation systems.

The rest of the paper is organized as follows. Section 2 summarizes related work. Section 3 describes the dataset. Section 4 shows the motivation of fine-grained travel time measurement, followed by design details in Section 5. Section 6 presents the measurement results. Section 7 implements a delay time prediction application. Section 8 discusses the lessons learned, insights, and limitations. Finally, Section 9 concludes the paper.

2 RELATED WORK

Travel time analysis has been investigated by considerable studies because of its importance to people's daily life. We summarize existing works in Table 1 with a two-dimension taxonomy: (i) transportation modality, i.e., single modality or multiple modalities; (ii) granularity, i.e., with or without travel time decomposition.

Table 1. Travel Time Measurement Survey

Categories		Granularity	
		Coarse-Grained	Fine-Grained
Transportation Modality	Single	[4] [9] [10] [13] [23] [26] [27] [40]	[5] [8] [11] [12] [39]
	Multiple	[30] [32] [34] [20] [29]	*MAC*

2.1 Studies on Single Modality

Most existing studies on travel time estimation are based on single modality due to the isolation of transportation systems, such as subways, buses, and taxis.

(i) Coarse-grained travel time: Given the task of estimating travel time, existing studies focus on the total travel time between the two locations in a city by bus [23], subway [9, 10], taxi [4, 13, 26, 40], or private vehicles [27]. The travel time estimation models are carefully designed for a specific transportation system or modality. The generalization is not investigated due to system isolation as well as data accessibility.

(ii) Fine-grained travel time: The actual travel time is not only the travel time in the transportation system such as between getting into the bus and getting off the bus. Instead, it is decomposed into different stages according to travel patterns. For example, the travel time is decomposed into waiting time in a bus station and riding time in a bus [8, 39] in bus systems. Similarly, the travel time is mostly divided into waiting time in a pickup location and the riding time in a taxi [5, 12]. For the subway system, the travel time is mainly divided into waiting time, riding time and transfer time [11].

2.2 Studies on Multiple Modalities

Studies involved with multiple modalities either integrate multiple mobile systems or investigate the differences and similarities between different mobile systems.

(i) Coarse-grained travel time: Due to the limitation of multi-modal data access, little work has been focused on coarse-grained travel time for multiple modalities, i.e., riding time from one location to another location based

on different systems. Some of the representative studies in this category include a multi-view learning model to explore human mobility using transportation and cellphone data [32], and an integration model to infer real-time traffic speeds with multi-source large-scale infrastructure data [34].
(ii) Fined-grained travel time: To the best of our knowledge, we are the first to investigate the fine-grained travel time pattern with multiple modalities. Building upon four transportation systems, our study aims to provide a comprehensive analysis on the fine-grained travel time pattern.

3 DATASETS

3.1 Investigated Transportation Systems

We are part of a team working under Shenzhen Smart City Initiative, through which we have access to four transportation fleets for our analysis. The details about the data are given in Table 2.

Table 2. Four Fleets

	Subway	**Bus**
Daily Data Size	2.23GB	6.5GB
# of Passengers	4 million	5 million
# of Daily Records	9 million	44 million
Format	Card ID Date&Time Station ID IN/OUT Fare Collection Machine ID	Plate Date&Time Stop ID GPS&Speed
	Personal Vehicle	**Taxi**
Daily Data Size	11 GB	6 GB
# of Vehicles	10,043	15,457
# of Daily Records	240 million	66 million
Format	Device ID Date&Time GPS&Speed	Plate Date&Time free/occupied GPS&Speed

- **Subway**: An automated fare collection (AFC) system is utilized to collect passenger's trip origins and destinations in the subway system when passengers tap in and tap out of subway stations. The data record includes *passenger id, tap-in station, tap-in time, tap-out station, tap-out time, two fare collection machine ID (IN&OUT).*
- **Bus**: There are two data sources in the bus system: The first data source is an automated fare collection (AFC) system, which records the trip details of bus passengers with smartcard swiping. The bus AFC data record includes *passenger id, tap-on station, tap-on time, tap-out station, tap-out time, bus id, bus route.* The second data source is an onboard bus GPS tracker, which has two components: a GPS recorder and a communication component. The GPS recorder logs the current bus GPS and time information, and the communication component sends the data to a cloud server. The record collected by the GPS tracker includes *id, longitude, latitude, time.* The accessed bus fleet data include 976 bus lines and 13 thousand buses. The bus fleet has a regular pattern due to their operating routes.
- **Taxi**: For the taxi data, in addition to the on-board GPS data, one more bit is recorded to indicate if the taxi is occupied by passengers or not. Therefore, one taxi record includes *id, longitude, latitude, time, occupied or not.* The taxi fleet in Shenzhen has 14 thousand taxis generating one status record every 30 seconds.
- **Private Vehicle**: We access Private Vehicle (PV) GPS trace data recorded by an onboard GPS device, which is installed by an insurance company for pay-as-you-go insurance programs. The GPS location is sent to the insurance company. The PV GPS records are similar to bus GPS records, containing *id, longitude, latitude, time.* We have access to this private vehicle network with more than 293 thousand vehicles, among which 10 thousand vehicles are in Shenzhen.

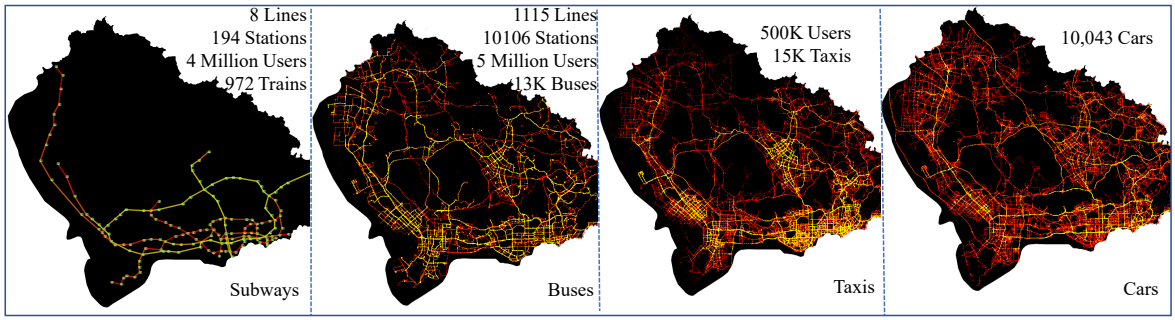

Fig. 1. Transportation Record Visualization

To illustrate the data distribution difference in the four transportation systems, Fig. 1 gives a heatmap visualization of these four transportation systems based on one-day data. We find some unique travel patterns in each transportation system. The taxi system covers most areas of the city. The bus system covers the main roads of the city. In the subway system, most of the congestion happens in the station. For the private vehicle system, it has similar travel coverage as the taxi system.

3.2 Anomalies

We list the investigated anomalies in two categories, i.e., expected and unexpected anomalies, in Table 3.

Table 3. Anomaly Categorization

Anomaly Category	Anomaly Examples
Expected	concerts, musical festivals, college entrance exams, subway fire drill, sports
Unexpected	station entrance of stagnant water, subway delay, facility malfunction, accidents

3.2.1 Unexpected Anomalies. Transportation accidents are one of the major unexpected factors that affect travel time and passengers' travel behaviors. However, accident datasets cannot be collected through city infrastructures directly with highly accurate spatial and temporal information. To investigate its impact on fine-grained travel time, we collect transportation accidents in the following procedures.
(i) *Retrieval.* We firstly build a web spider to extract and obtain all the posts from official accounts of the Shenzhen Transportation Police from *Sina Weibo*, i.e., the Chinese version of *Twitter*. We filter the raw data with the keywords *road condition* or *accident*. By this approach, we focus on the relevant information and use the texts posted on the website for anomaly analysis.

Table 4. Description of Accident Data Set

Description keywords	Anomaly Level
scrape, small cars, slow	1
bus, van, injured, several cars	2
fire, dead, construction	3

(ii) *Structuralizing.* We extract the useful information from the texts we have crawled in the previous step. We apply regular expressions to obtain the time and location of the accidents. To further quantify the extent of anomaly, we classify them into different *anomaly levels* according to the keywords in the description provided. For example, if the keywords such as "small cars" and scrape" are mentioned, we consider it as a low level of severity. The details of anomaly level and its corresponding keywords are provided in Table 4.
(iii) *Matching locations. Geocoding* is the process of converting addresses (e.g., "1600 Amphitheatre Parkway, Mountain View, CA" or "Intersection of 5th street and 62nd street in NYC") into geographic coordinates (e.g.,

latitude 37.423021 and longitude -122.083739). *Geocoding* is conducted with two steps, i.e., *Entity Extraction* and *Location Matching* with Google Map API [17].
(iv) *Data cleaning*. We clean out some unreasonable and unrealistic data such as the ones with coordinates out of the region of Shenzhen. Finally, we obtain Table 5 to describe every item in the data set of unexpected anomaly.

Table 5. Sample of Unexpected Anomaly Data Set

Events	Description	Time	Location	Latitude	Longitude	Level
A	many cars collided	2016-06-01 14:42	BeiHuan Ave, Shahe West Overpass	22.56829	113.95432	2
B	two cars scraped	2016-06-01 08:55	NanPing Exp, TangLangShan Road	22.56748	113.99177	1
C	a van turned over	2016-06-04 14:25	ShenShan Highway	22.53988	114.01955	3

3.2.2 Expected Anomalies. Expected social events such as Marathon, concerts and music festivals have a large impact on passengers' travel behavior. We collect these social events by crawling and analyzing online news. The crawling procedures are similar to the irregular anomaly extraction, which includes *event retrieval*, *event structuralizing*, *location matching* and *data cleaning*. We show the samples of the data set of regular anomaly in Table 6.

Table 6. Sample of Expected Anomaly Data Set

Events	Description	Time	Location	Latitude	Longitude	Level
A	Marathon	2016-12-04 8:00	Baoan People's Government	22.553623	113.884333	3
B	Midi Music Festival	2016-01-01	Universiade Shenzhen Gymnasium	22.693345	114.219014	2
C	Emergency Drill	2016-06-26 1:00	Futian Subway Station	22.541603	114.052626	1

4 MOTIVATION

4.1 Fined-grained Travel Time

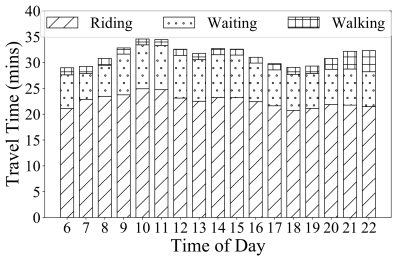

Fig. 2. Fine-grained Travel Time

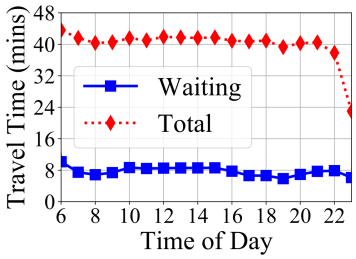

Fig. 3. Waiting Time

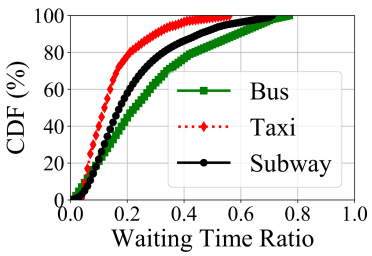

Fig. 4. Waiting Time Ratio

We compare the fined-grained travel time, i.e., riding, waiting, and walking, with coarse-grained travel time, i.e., total travel time or riding time. First, we study the average waiting time, walking time and riding time that subway passengers spend between two major subway stations with high traffic, i.e., central station. Fig. 2 plots the fine-grained travel time distribution between these two stations in the subway system. We find that even though passengers spend a large portion of travel time on riding, the waiting time and walking time are not negligible and have higher fluctuation. Fig. 3 compares the waiting time with the total travel time in all subway trips. On average, the waiting time accounts for 16% of the total travel time in the subway system. Moreover, we compare the waiting time with the total travel time in the bus, taxi and subway systems. Fig. 4 shows the

CDF of the ratio of waiting time in the total travel time in all trips. We find that the waiting time accounts for more than 20% of the total travel time in 20% of the trips in the taxi system. In the bus system, more than 20% of the travel time is spent on waiting in 57% of all trips. In the subway system, passengers spend more than 20% of the travel time on waiting for trains in 40% of all trips. Compared with the coarse-grained travel time, the fined-grained travel time accurately describes passengers' travel time in different stages. The details of travel time decomposition will be provided in Section 5.4.

4.2 Multiple Transportation Systems

Passengers with travel demand in a city dynamically choose one or multiple of the transportation modalities, e.g., bus, subway, taxi, and private vehicle, based on travel purposes and traffic conditions. For instance, in our analysis, we find that commuters traveling between home and work locations prefer to take public transportation systems or drive their own cars instead of riding taxis. In contrast, visitors mostly take taxis at airports or train stations. We show the travel demand (i.e., the number of passengers) and the demand trend (i.e., the normalized travel demand) in the four transportation systems in Fig. 5 and Fig. 6. We find some obvious differences among these four systems in terms of travel demands. The two public transportation systems have two peaks of travel demand during rush hours. The taxi demand increases significantly in the evening due to different travel purposes. Therefore, a single transportation system is not the representative for travel behaviors in a city due to its biased mobility patterns.

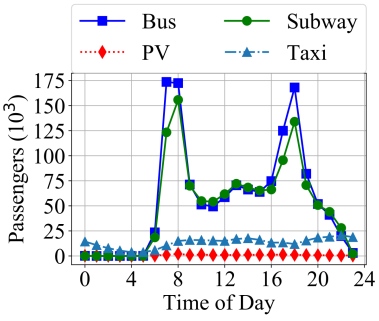

Fig. 5. Travel Demand

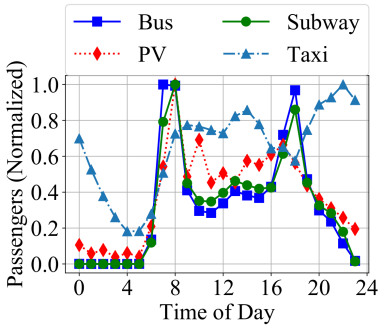

Fig. 6. Demand Trend

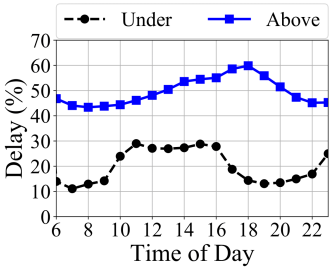

Fig. 7. Impact of the Storm

4.3 Anomaly Events

Anomaly events, such as social events and transportation accidents, affect the traffic flow, which leads to travel delay. For example, the severe tropical storm Nida in 2016 has caused severe travel delay. We quantify the impact of the storm on the travel time by the increased travel delay, which is calculated as the relative difference between travel time. We study its impact of travel time on both above-ground and underground transportation systems. Fig. 7 shows the increased travel delay during the storm. The storm significantly increases the travel time on above-ground systems, i.e., buses, taxis, and private vehicles by around 40% to 70%. It also affects the service of several subway lines and stations and leads to travel delay in the underground subway system as well.

4.4 Summary

In summary, we investigate the features of fine-grained travel time and anomalies. Even if passengers spend a large portion of time on riding during a trip, the waiting time is not negligible as it makes up around 20% of the total travel time. Due to mobility difference in different transportation systems, a single transportation system is not representative of all passengers' travel behavior in the city. This motivates us to measure the impact of

anomalies on fine-grained travel time, i.e., waiting time and riding time on four different transportation systems to better understand their impacts on urban mobility.

5 METHODOLOGY

This section describes how we decompose trips into travel components to better understand urban mobility. We first describe passengers' behaviors with different transportation modalities. Second, we elaborate on travel time decomposition to infer fine-grained travel time components.

5.1 Background: Understanding Passenger Behaviors

In most cities, the major travel modalities include subways, buses, taxis and private vehicles. In order to decompose the total travel time into fine-grained time components, we first aim to understand the travel behaviors in passengers using different transportation systems.

(i) **Taxi**. In the taxi system, passengers walk from an origin to the pickup location and wait for the next available taxi. Next, passengers take a taxi to a drop-off location, after which they walk to their final destination. Most of the travel time is spent on *waiting* and *riding*.

(ii) **Bus**. If passengers choose to take a bus, similar to the taxi, passengers first walk to a bus station. Then passengers wait in the bus station for the next available bus to their destination station. Similar to taxi passengers, bus passengers spend most of the travel time on *waiting* and *riding* as well.

(iii) **Private Vehicle**. For private vehicle passengers, since there is generally no waiting time for pickup, most of the travel time is spent on *riding*.

(iv) **Subway**. For the subway system, passengers' behavior is decomposed into *walking*, *waiting* and *riding*. Starting from an automated fare collection machine in a station, passengers *walk* to a train waiting platform. Next, passengers wait in the platform for the next available train. In the third phase, passengers take the train to the destination station and then walk to the tap-out machine to exit the subway system. Subway passengers spend most of the travel time in *walking*, *waiting* and *riding*.

Summary. By analyzing passengers' behavior, we specifically investigate (i) *riding time* for personal vehicles, (ii) *waiting time* and *riding time* for taxis and buses, and (iii) *waiting time* and *riding time* for subway systems. The analysis does not include in-station *walking time* since the dynamics of in-station walking time is negligible.

5.2 Terminologies

In this subsection, we describe the terminologies and their abbreviations in the travel time decomposition.

Transportation System. We use four initial characters to represent the four investigated transportation systems. Specifically, we use $\mathcal{B}$ to represent the city bus system, $\mathcal{S}$ for the city subway system, $\mathcal{T}$ for the taxi system and $\mathcal{P}$ for personal vehicles. We categorize the four systems into above-ground group $\mathcal{A}$ and underground group $\mathcal{U}$, since above-ground systems share the same mobility pattern and road traffic in travel time components.

Travel Time Components. The total travel time consists of three travel time components, i.e., walking, waiting and riding, based on semantics of passengers' behavior. A travel time component is described as $\tau^{\mathcal{M}}_{status}(s, t)$ where $\mathcal{M}$ and *status* are two global parameters to describe the transportation modality/system and the status of passengers (e.g. walking or waiting). The two spatiotemporal parameters s and t are location and time respectively. For travel time components involving two locations, e.g., riding time with an origin and a destination, we use two parameters, i.e., $\tau^{\mathcal{S}}_{status}(s_1, s_2, t)$, to present the locations. In addition, an individual-level travel time component is denoted as $\tau^{\mathcal{S}}_{status}(s, t, i)$, where i represents the id of a particular passenger.

Anomaly Event. Similar to a travel time component, an anomaly event $\mathcal{E}^{level}_{category}(s, t)$ is associated with two global parameters and two spatiotemporal parameters where the two global parameters $level$ and $category$ are the extent of the anomaly event defined in the dataset section, and the category of the event (i.e., expected or

unexpected event). The two spatiotemporal parameters s and t describe the location and time of the anomaly event.

5.3 Travel Time Decomposition: Aboveground

Given city-scale mobility data in four transportation systems, we describe how to infer the fine-grained travel time components in a specific transportation system. Specifically, we elaborate the inference of (i) the waiting time in taxi, bus, and subway systems, (ii) the riding time in the four systems, and (iii) the in-station walking time in the subway system.

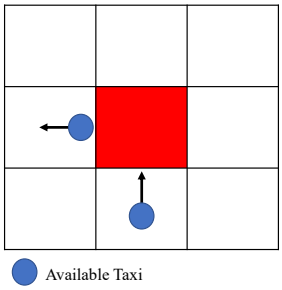

Fig. 8. Waiting Time Inference

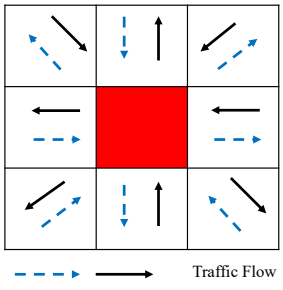

Fig. 9. Riding Time Inference

5.3.1 Waiting Time $\tau^{\mathcal{A}}_{waiting}$. We use the taxi system as an example to illustrate the idea of waiting time inference. As shown in Fig. 8, given taxi traces passing through location s, the upper bound of the waiting time is inferred by the time difference of two successive available taxis which accommodate at least one passenger. For buses, the waiting time is inferred by the time difference of two successive buses from the same route. Therefore, we divide the city into $50m \times 50m$ grids, which leads to 1800×920 grids in the city. In addition, the waiting time is affected by the number of passengers in the waiting queue. For example, if there are two passengers waiting in the same location for taxis at the same time without ride sharing, the statistical waiting time is the average waiting time of the two passengers. For the first passenger, the waiting time is obtained through the waiting time estimation of the next available taxi. For the second passenger, in addition to the waiting time of the next available taxi, the extra time is the queuing time that equals the waiting time of the first passenger. This total queuing time is determined by the travel demand which is estimated through the historical records of the taxi data.

$$\tau^{\mathcal{A}}_{waiting}(s,t) = \frac{\sum_{n=1}^{\mathcal{D}(s,t)}(r_{i+n}.time - r_i.time)}{\mathcal{D}(s,t)}; \quad (1)$$
$$s = r_i.location; \;\; r_i.time \; in \; t \; time \; slot;$$

For location s, e.g., a grid for taxis and a station for buses, we estimate the waiting time $\tau^{\mathcal{A}}_{waiting}(s,t)$ at time t in Equation (1) where r_{i+n} is next n available vehicles after vehicle i and $\mathcal{D}(s,t)$ is the travel demand at location s and time t. We model the demand by the number of passengers in the historical data, which is captured by the taxi status changing from 0 (empty) to 1 (occupied). Since it is almost impossible to infer the large-scale exact waiting time of individuals, our inference model focuses on the upper bound of individual roadside waiting time based on taxi GPS locations.

5.3.2 Riding Time $\tau^{\mathcal{A}}_{riding}$. The riding time inference from historical data has been extensively investigated in the previous works [26] [39]. Instead of estimating travel time on road segments, we infer the riding time in grids before and after an anomaly happens. We first investigate the riding time from the traffic flows that have to pass

the grid in the normal cases. Then, to compare the travel time difference before and after an anomaly event in the central red area, we compare the average riding time from eight directions, i.e., from the top to the bottom, from the left to the right, from the upper left to the bottom right, from the upper right to the bottom left and the reverse directions, as shown in Fig. 9.

5.4 Travel Time Decomposition: Underground

Different from above-ground systems, which mainly use GPS devices to track vehicle locations, underground systems depend on stationary devices, e.g., automated fare collection systems, to track passenger flows. For each trip, the total travel time is captured by autonomous fare collection machines, which record time and location when passengers tap in or tap out the subway system, i.e., the origin and destination stations. Therefore, the total travel time of taking the subway is the time difference between the tap-in time and tap-out time of a passenger. With travel flows from an origin station s_o to a destination station s_d, the total travel time between the two locations is $\tau^U_{total}(s_o, s_d, t, i)$. As described in passenger behavior analyses, the total travel time is decomposed into walking, waiting and riding time components, as shown in Equation (2).

$$\begin{aligned} \tau^{\mathcal{S}}_{total}(s_o, s_d, t) = \tau^{\mathcal{S}}_{walking}(s_o, t) + \tau^{\mathcal{S}}_{waiting}(s_o, t_1) + \tau^{\mathcal{S}}_{riding}(s_o, s_d, t_2) + \tau^{\mathcal{S}}_{walking}(s_d, t_3); \\ t_1 = t + \tau^{\mathcal{S}}_{walking}(s_o, t);\ t_2 = t_1 + \tau^{\mathcal{S}}_{waiting}(s_o, t_1);\ t_3 = t_2 + \tau^{\mathcal{S}}_{riding}(s_o, s_d, t_2); \end{aligned} \tag{2}$$

5.4.1 Waiting Time $\tau^{\mathcal{S}}_{waiting}$. To illustrate the inference of the waiting time for a passenger i in a station s_o, we first define the *fluent travel time*. We infer the fluent travel time with the following observation in subway systems. Most of the trips include four stages: i) the walking time in the original station, ii) the waiting time in the original station, iii) the riding time from original station to the destination station, and iv) the waiting time in the destination station. Since the riding time following the subway schedule is constant and the difference of walking time among passengers is negligible, the minimal travel time occurs when "lucky" passengers catch the subway without waiting. We define the fluent travel time as the travel time without waiting which is inferred based on those "lucky" passengers. Therefore, the fluent travel time is the travel time between two stations without the waiting time, which is estimated by the minimum travel time between two stations. The waiting time is the time difference between the total travel time and the fluent travel time. The inference is described in Equation (3).

$$\begin{aligned} \tau^{\mathcal{S}}_{fluent}(s_o, s_d, t) = \min_i \tau^{\mathcal{S}}_{total}(s_o, s_d, t, i); \\ \tau^{\mathcal{S}}_{waiting}(s_o, s_d, t, i) = \tau^{\mathcal{S}}_{total}(s_o, s_d, t, i) - \tau^{\mathcal{S}}_{fluent}(s_o, s_d, t); \end{aligned} \tag{3}$$

5.4.2 Walking Time $\tau^{\mathcal{S}}_{walking}$. We use an example to show how to extract the walking time in a station, i.e., station B, in Fig. 10 where the dash arrow represents the walking time in the station and the solid line represents the riding time between two stations.

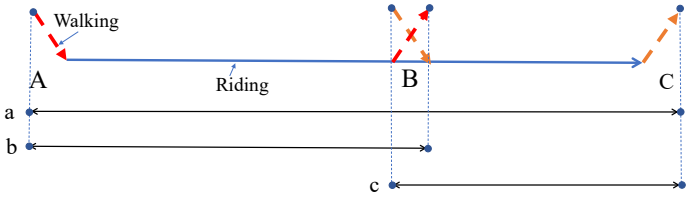

Fig. 10. Subway Passenger Behavior

Given three fluent trips (i.e., trips without waiting time), trip a from station A to station C, trip b from station A to station B, trip c from station B to station C, B is the intermediate station of A and C. The fluent travel time in the three trips is decomposed into walking and riding time, as shown in Equation (4).

$$\begin{aligned}
\tau^{\mathcal{S}}_{fluent}(A, C, t') &= \tau^{\mathcal{S}}_{walking}(A, t') + \tau^{\mathcal{S}}_{riding}(A, C) + \tau^{\mathcal{S}}_{walking}(C, t'') \\
\tau^{\mathcal{S}}_{fluent}(A, B, t') &= \tau^{\mathcal{S}}_{walking}(A, t') + \tau^{\mathcal{S}}_{riding}(A, B) + \tau^{\mathcal{S}}_{walking}(B, t) \\
\tau^{\mathcal{S}}_{fluent}(B, C, t) &= \tau^{\mathcal{S}}_{walking}(B, t) + \tau^{\mathcal{S}}_{riding}(B, C) + \tau^{\mathcal{S}}_{walking}(C, t'')
\end{aligned} \tag{4}$$

The riding time of trip a is the summation of the riding time of trip b and the riding time of trip c.

$$\tau^{\mathcal{S}}_{riding}(A, C) = \tau^{\mathcal{S}}_{riding}(A, B) + \tau^{\mathcal{S}}_{riding}(B, C) \tag{5}$$

In Equation (4), we find the second equation plus the third equation minus the first equation, we can remove the $\tau^{\mathcal{S}}_{walking}(C, t'')$. After applying Equation (5), the riding time will be removed. Therefore, the walking time of station B is inferred by the fluent travel time difference of the three trips, which is described in Equation (6).

$$\begin{aligned}
\tau^{\mathcal{S}}_{walking}(B, t) = \frac{1}{2}[\tau^{\mathcal{S}}_{fluent}(A, B, t') + \tau^{\mathcal{S}}_{fluent}(B, C, t) - \tau^{\mathcal{S}}_{fluent}(A, C, t')]; \\
t = t' + \tau^{\mathcal{S}}_{fluent}(A, B, t');
\end{aligned} \tag{6}$$

5.4.3 Riding Time τ^{S}_{riding}. Given the *waiting time* and *walking time*, according to Equation (2), the *riding time* inference is straightforward, which is the difference of the total travel time and other time components, as shown in Equation (7).

$$\tau^{\mathcal{S}}_{riding}(s_o, s_d, t_2) = \tau^{\mathcal{S}}_{total}(s_o, s_d, t) - \tau^{\mathcal{S}}_{walking}(s_o, t) - \tau^{\mathcal{S}}_{waiting}(s_o, t_1) - \tau^{\mathcal{S}}_{walking}(s_d, t_3); \tag{7}$$

Summary. With the decomposition method, we infer the walking time and waiting time in subway stations, as well as the riding time between subway stations. In the implementation, we find there are multiple entrances in large stations, e.g., 12 entrances in Futian station which is adjacent to the city train station, and the walking time difference among different entrances is not negligible. Fortunately, since there are always high traffic demands in large stations, it provides enough observations in each entrance. On the other hand, different entrances can be identified by the id of fare collection machine in the dataset. Therefore, we identify the large stations by a threshold on the number of entrances, i.e., fare collection machines. Instead of modeling travel time in the station level, we infer the walking time, riding time and waiting time on entrances in the large stations. When calculating the average waiting time in the station or riding time between stations, we aggregate inferred time from multiple entrances.

5.5 Measurement Method

Anomaly Measurement: We investigate the impact of an anomaly at location s and time t on a travel time component $\tau^{\mathcal{M}}_{status}(s, t)$ in multiple transportation systems in a city. With the anomaly measurement task, we are able to study the impact of an anomaly event $\mathcal{E}$ on the travel time components in multiple transportation systems at the city level. As a result, it enables us to compare the robustness of different transportation systems under abnormal conditions at extreme fine-grained travel patterns, which hopefully will provide insights for urban planning [3], navigation [18], travel planning [28] and anomaly detection systems [1].

Metrics: In addition to the direct comparison of travel time components, we use delay time in Equation (8) since travel time components differ at different spatiotemporal dimensions, and the delay time is defined as the time difference between the time that passengers spent under the impact of anomalies and the time that passengers spent without anomalies.

$$d^{\mathcal{M}}_{status}(s,t|\mathcal{E}) = \tau^{\mathcal{M}}_{status}(s,t|\mathcal{E}) - \tau^{\mathcal{M}}_{status}(s,t) \tag{8}$$

System-level measurement. To understand the impact of anomalies, for every system, we study $d^{\mathcal{M}}_{status}$ under different types of anomalies with certain degrees. For an event $\mathcal{E}^{level}_{category}(s,t)$, we compare the anomaly events with different *category* and study their impact on travel time in individual systems.

Inter-system comparison. To compare the robustness of different transportation systems under the impact of anomalies, we study the impact of anomalies of the same category on different transportation systems.

6 MEASUREMENT RESULTS AND ANALYSIS

In this section, we first briefly evaluate our travel time decomposition methods with case studies since it is the foundation of the anomaly measurement. Second, we study the fine-grained travel time patterns in the four systems. Third, we investigate the impact of anomalies on the travel time by the delay time.

Since the riding time estimation is a statistical aggregation of observations and the waiting time is inferred from the data, we validate the waiting time inference by case studies. We conduct case studies by videotaping the passenger flow in both subway stations and bus stations to validate the $\tau^{\mathcal{M}}_{waiting}$ for both underground and above-ground systems as shown in Fig. 11 and Fig. 12. We collect videos recording passengers' waiting behaviors in the subway stations and the bus stations in both peak hours (i.e., morning rush hours and evening rush hours) and non-peak hours (i.e., regular time) in one week, which covers 1181 subway passengers and 677 bus passengers. For the subway system, we count passengers' waiting time by recording the time they arrive at the station and

Fig. 11. Subway Passenger Behavior

Fig. 12. Bus Passenger Behavior

the time they get on the train. Due to the space limitation, we show the waiting time distribution during two peak hours (08:30-09:30 and 17:30-18:30) and one regular time (13:30-14:30) in Futian subway station in Fig. 13, Fig. 14, and Fig. 15, which cover the waiting behavior of 241 out of 1181 subway passengers in our dataset.

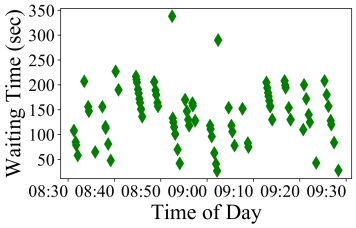

Fig. 13. Subway - Morning Peak Hour

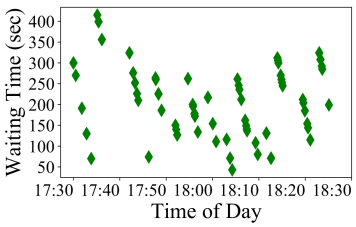

Fig. 14. Subway - Evening Peak Hour

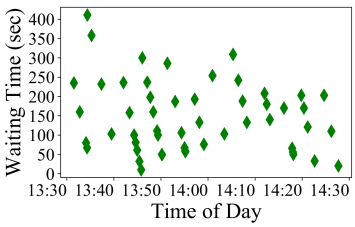

Fig. 15. Subway - Regular Time

Fig. 16 shows the inferred waiting time from subway transaction records. We first use a noise reduction by a distance constraint between points to remove noises. Second, we divide the points into clusters by a clustering

algorithm, e.g., DBSCAN. In the third step, we apply an in-cluster regression algorithm on the inferred waiting time. We show the process in Fig. 17. We then compare the observations with the regression center at the same time. The average root mean square error is 0.829 minutes in the case study, which shows the inference can capture passengers' waiting behaviors in the subway station. We use the waiting time in bus stations to validate

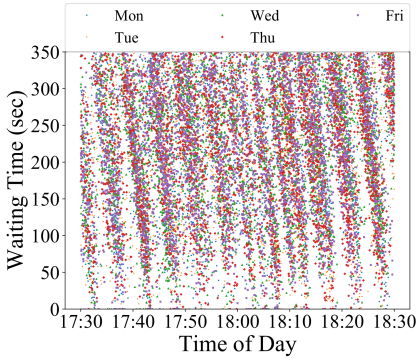

Fig. 16. Subway Inference

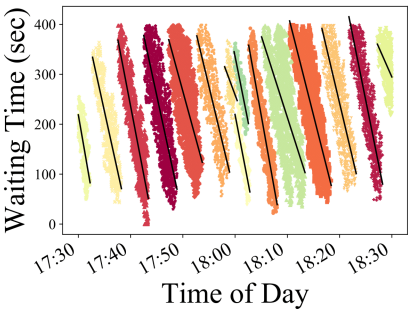

Fig. 17. Subway Inference Regression

the aboveground waiting time inference. Fig. 18 to Fig. 20 shows the waiting time of 138 passengers in Futian bus station in both peak hours and regular time. The waiting time distribution in the subway system is different from that in the bus system, i.e., the waiting time is more regular in the subway system compared with the bus system. The difference is caused by the waiting patterns and schedules of the two systems. In the subway system, subway passengers wait for the same train while bus passengers wait for different buses. Therefore, the waiting time in the subway system is linearly related with the time when the passengers start the waiting while the waiting time in the bus system is additionally determined by a specific bus arrival time. Since the above-ground waiting time inference estimates the upper bound of the waiting time in the bus system and the taxi system, we use the precision score as the accuracy metrics. If the passengers' waiting time is in the estimated bounds, the true positive will increase by 1. In this way, the precision score is 96.8% for the bus system in the case study.

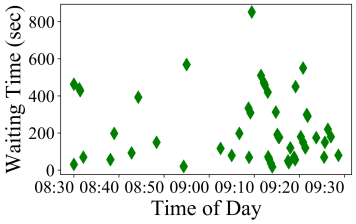

Fig. 18. Bus - Morning Peak Hour

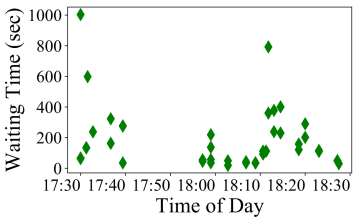

Fig. 19. Bus - Evening Peak Hour

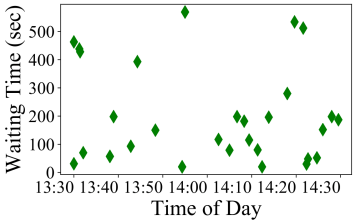

Fig. 20. Bus - Regular Time

6.1 Waiting Time Patterns

6.1.1 Above-ground System. Based on our waiting time inference of above-ground transportation systems, we study the the waiting time of the three transportation systems, i.e., taxi, bus and subway. We compare the aboveground waiting time distribution of weekdays and weekends in Fig. 21. Comparing the weekday pattern with weekend pattern, we find that the main difference of waiting time is located at around 5pm, which is the peak hour in the afternoon.

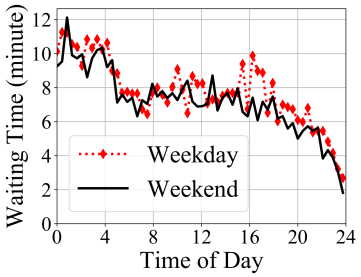

Fig. 21. Aboveground Waiting Time

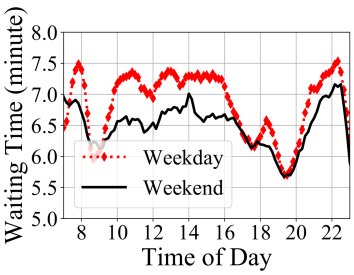

Fig. 22. Underground Waiting Time

To study the difference in the spatial dimension, we visualize the waiting time distribution of the peak on the spatial dimension in Fig. 23 and Fig. 24 for taxi systems. We find that the waiting time is longer in the downtown area and the transportation junctions such as the airport and the train station. Similar pattern has been found in the bus system as shown in Fig. 25 and Fig. 26. The waiting time is higher on peak hours in the weekdays in the bus system and around 18.9% of the buses cannot follow the regular time table schedules.

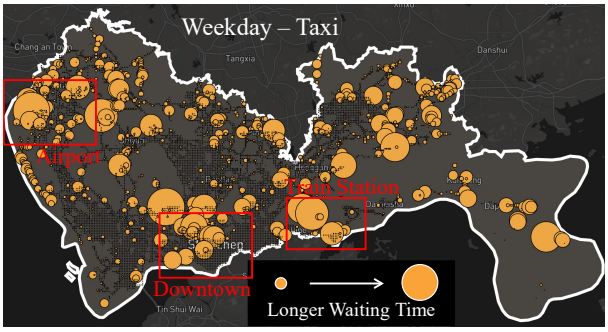

Fig. 23. Taxi Waiting - Weekdays

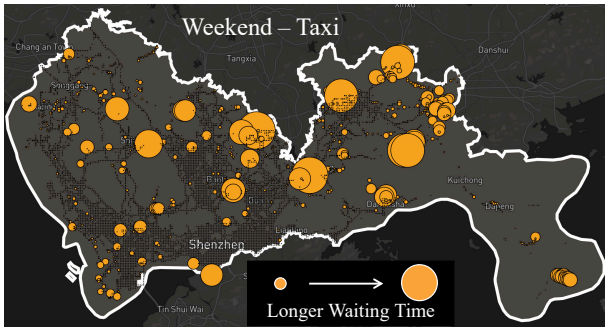

Fig. 24. Taxi Waiting - Weekends

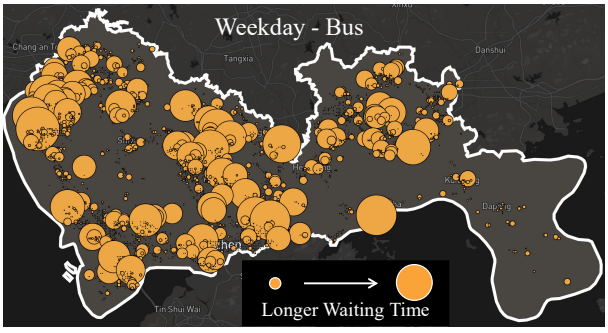

Fig. 25. Bus Waiting - Weekdays

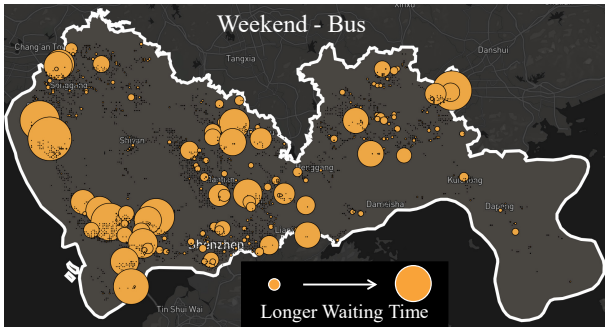

Fig. 26. Bus Waiting - Weekends

6.1.2 Underground System. Fig. 22 compares waiting times on weekdays and weekends. On weekdays, the waiting time is longer in non-peak hours during daytime and nights. However, in the two peak hours, the waiting time is lower compared with non-peak time. The reason is that subway operators send more trains in the peak hours, which decreases the time interval between two consecutive trains in a subway station. Besides, the waiting time difference is small between weekdays and weekends during peak hours. The reason is that passengers in peak hours are most commuters, who are sensitive with time. Those passengers take the train although the trains are highly loaded. However, during the non-peak hours, since there are more passengers on weekdays, unlike

commuters, those passengers prefer to wait for next available trains if the current one is highly loaded. Therefore, it increases the waiting time compared with weekends.

Fig. 27 and Fig. 28 present the waiting time distribution on the spatial dimension. Although the spatial distribution is similar between weekdays and weekends, we find higher waiting time in the downtown areas during weekdays. Fig. 29 shows the weekly pattern of the waiting time distribution. Motivated by the cellphone

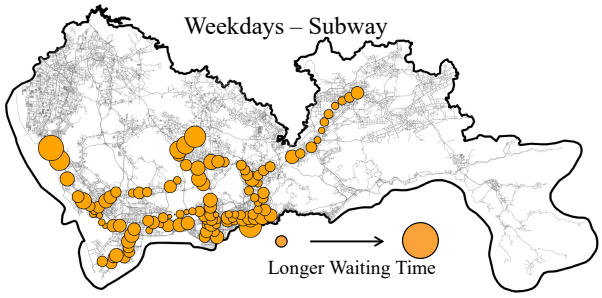

Fig. 27. Subway Waiting - Weekdays

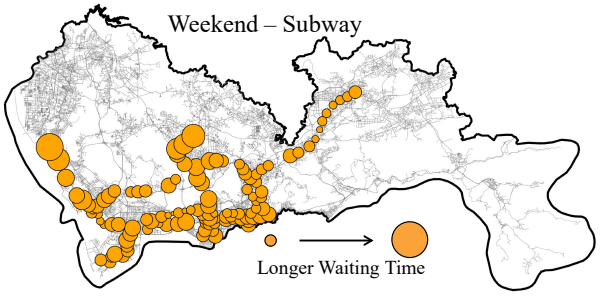

Fig. 28. Subway Waiting - Weekends

tower clustering by traffic patterns [25], we apply k-mean clustering on the normalized weekly patterns. According to the tuning of parameters, the best k is 4. The result is shown in Fig. 30, in which each line indicates different functions of regions where the stations locate.

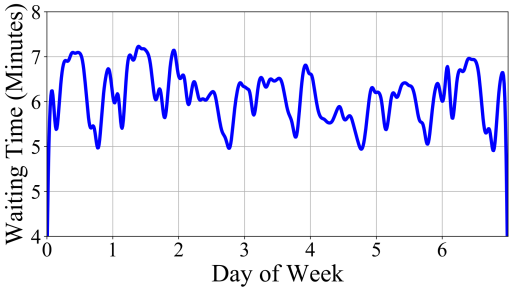

Fig. 29. Subway Average Waiting Time

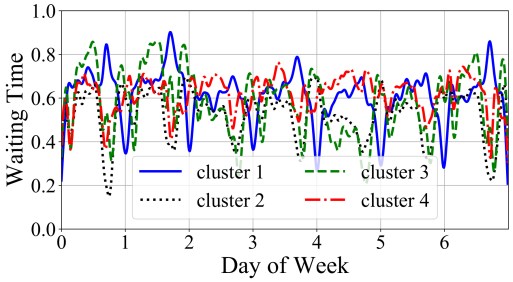

Fig. 30. Subway Waiting Time Patterns

6.2 Riding Time Patterns

Since the three above-ground systems present very similar riding time patterns due to the shared road networks, we present their average riding time distribution in Fig. 31. The riding time in the aboveground systems is inferred by the average time that aboveground vehicles spend in passing a grid ($50m \times 50m$). In weekdays, especially during the peak hours, the riding time increases dramatically due to poor traffic conditions. Moreover, the peak hours shift from 8am and 6pm on weekdays to 11am and 5pm in weekends, which are the time for lunch or dinner. For the subway system, we find that the riding time on weekdays and weekends between the same origin and destination stay almost unchanged. Therefore, we compare the riding time of individual passengers to study their travel behaviors. Fig. 32 shows the cumulative distribution where passengers take much more riding time during weekends. The reason is that a large portion of trips on weekdays are commuting between home, work and central business areas. Since passengers live around their work locations, the home-work distance is not far. In contrast, the trips during weekends are random trips, e.g., from home location to parks with families. Therefore, the distances of trips during weekends are longer compared with those on weekdays.

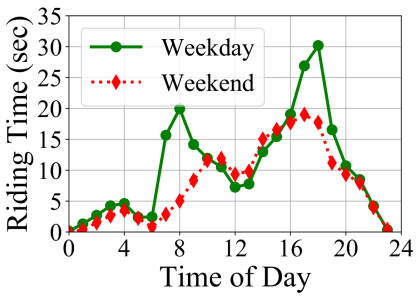

Fig. 31. Riding - Above

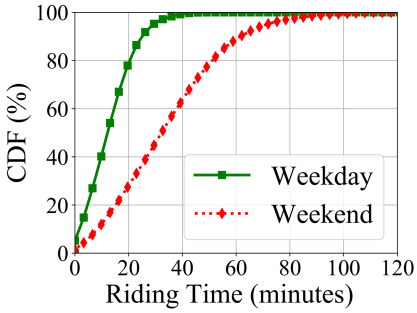

Fig. 32. Riding - Subway

6.3 Impact of Anomalies

We measure the overall impact of anomalies on the travel time components in the four transportation systems. We show the cumulative distribution of delay waiting time under the impact of two categories of anomalies in Fig. 33, Fig. 34 and Fig. 35. In general, the unexpected anomalies have a larger impact on the travel time components compared with the expected anomalies. Compared with the bus and taxi system, the subway system has lower delay time in expected anomalies but higher delay time in unexpected anomalies. The reason is that unexpected anomalies cause a large impact on the subway operation (e.g., closing subway lines or changing the schedule of subway trains).

The taxi system shows higher delay time in expected events compared with the bus system. In the analysis, we find that in expected anomalies, such as concerts, the increase of travel demand increases the waiting time. In the bus waiting time inference, we assume that buses accommodate all passengers. Therefore, the taxi delay time is caused by traffic delay and the availability of taxis while the bus delay time is caused by the traffic conditions. In terms of unexpected events, i.e., accidents, the taxi systems show a lower waiting time increase. One reason is that taxis are more agile with poor road conditions.

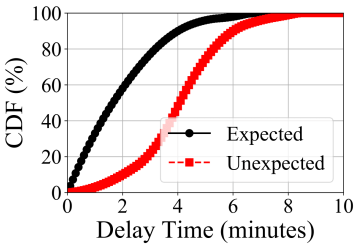

Fig. 33. Taxi Waiting

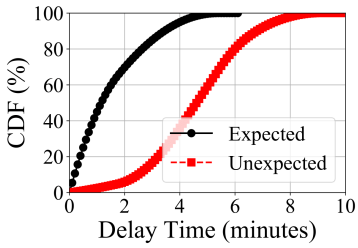

Fig. 34. Bus Waiting

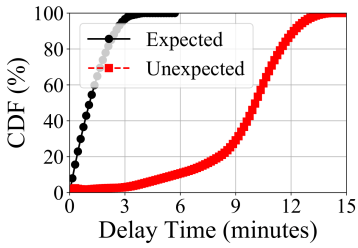

Fig. 35. Subway Waiting

We show the cumulative distribution of delay riding time under the impact of two categories of anomalies in Fig. 36, Fig. 37 and Fig. 38. We ignore the riding time analysis for the subway system since we find that the riding time is rarely affected by anomalies inside or close to the subway stations. As shown in Fig. 9, to capture the travel delay caused by the detours in the anomalies, we compare the travel time of eight directions of traffic flows between the normal cases and anomalies. Since the three above-ground systems share the same road infrastructures, the riding time changes are similar in the three transportation systems. Among the three transportation systems, the taxi system is most stable and the bus system is the least stable in terms of anomalies. The possible reason is that taxis choose detours while buses have a constant route, which may cover the regions where accidents happen. Compared with personal car drivers, the taxi drivers are more experienced. This leads to a lower delay in riding time in the taxi system.

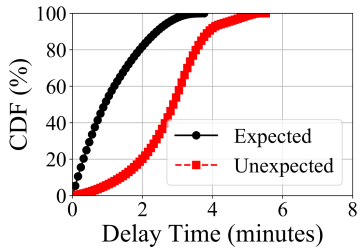

Fig. 36. Taxi Riding

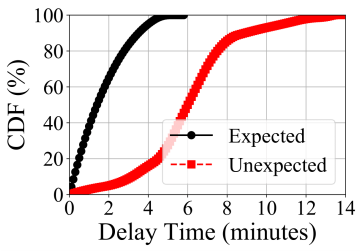

Fig. 37. Bus Riding

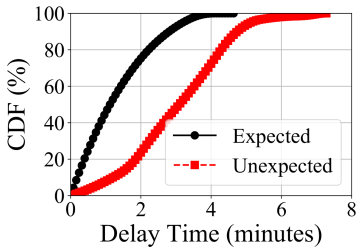

Fig. 38. PV Riding

7 APPLICATION: DELAY TIME PREDICTION

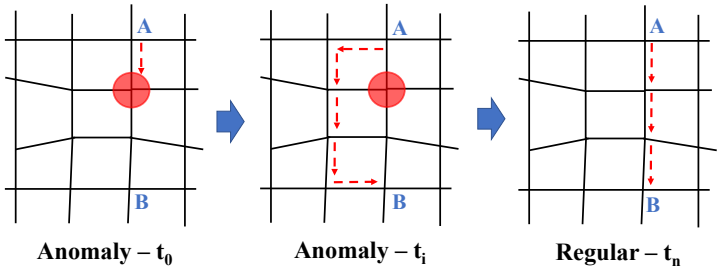

Fig. 39. Delay Time Prediction

In this section, we use the measurement results to predict delay time in anomalies. As shown in Fig. 39, given an observation that an anomaly $\mathcal{E}^{level}_{category}(s, t_0)$ happens at location s and time t_0, our target is to predict the delay time in the following n time slots $\{d(s, t_0), d(s, t_1), \cdots, d(s, t_{n-1})\}$ for a travel time component τ.

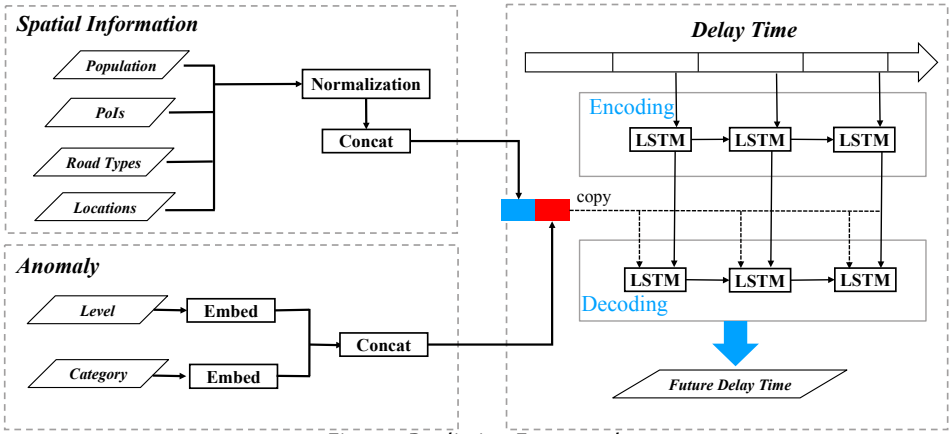

Fig. 40. Prediction Framework

7.1 Prediction Framework

As shown in Fig. 40, the prediction framework consists of three components: (i) an anomaly information feeder to feed anomaly features, (ii) a spatial information feeder to feed spatial features, and (iii) a long short-term memory

(LSTM) learning model that takes a sequence of delay time after the anomaly and predicts the future delay time. We describe the details of each component as follows.

(i) Anomaly Information. In the measurement study, we find that the impact of anomaly events differs with categories and levels at different time. Therefore, we use three features of the anomalies in our prediction including *anomaly category*, *level* and *time*. Since the three features are all categorical features, which cannot be directly used as the input for the learning model, we add an embedding layer before feeding them into the learning model.

(ii) Spatial Information. Figs. 23-26 have shown that the spatial dimension influences the travel time components significantly. We use different features related to the locations where the anomaly happens. Besides the geographic value of the location, the spatial features include *population*, *PoIs* (i.e., point of interests), and *road types*. The population information is extracted from the Worldpop data set [19], which provides population

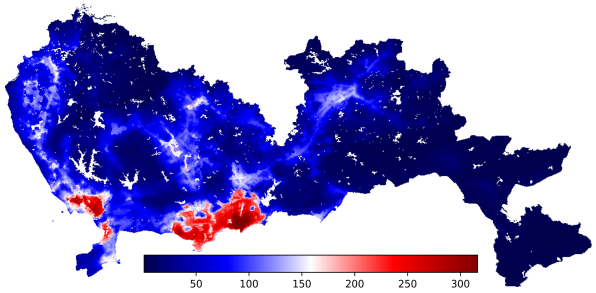

Fig. 41. Population Distribution

Fig. 42. Road Network

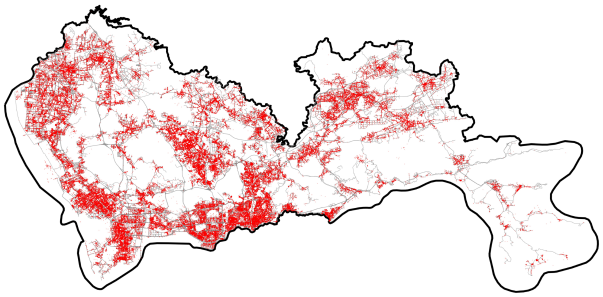

Fig. 43. PoI Distribution

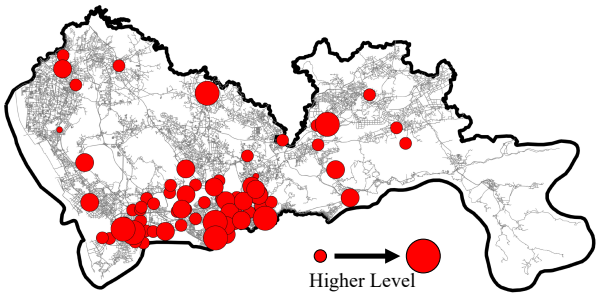

Fig. 44. Anomaly Distribution

distribution with high resolution, i.e., $100m \times 100m$. We map the population under the preset partition to our partition based on the intersection area of the two partitions [2]. We access the road networks and PoIs from an online service provider OpenStreetMap [6]. We categorize roads into four classes, i.e., main roads, secondary roads, path and highway, and the PoIs into five groups, i.e., educational, residential, office, recreation and transport, based on the labels in OpenStreetMap. We use the distribution of PoIs and road segments in regions as spatial features. We visualize the population distribution in Fig. 41 , the road network in Fig. 42, the PoI distribution in Fig. 43, and the anomaly distribution in Fig. 44. We find higher population on road segments and more PoIs in the central business district, i.e., the middle bottom area on the map, compared with other areas. The distribution of spatial contextual information, i.e., population, roads and PoIs , shows a positive correlation with the occurrence of anomalies on the spatial dimension. To reduce the impact of data scales, we apply minmax scaler on the numerical values of all regions for normalization.

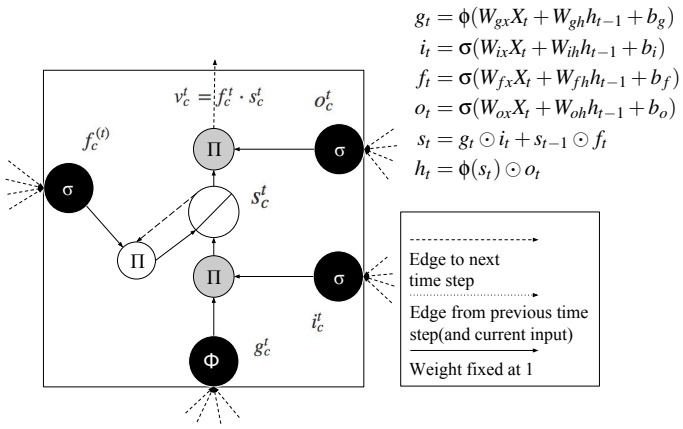

Fig. 45. Prediction Framework

(iii) Delay Time Learning. Due to the spatial-temporal nature of human mobility, the delay time distribution presents a high correlation with the spatial and temporal information. Recurrent Neural Network (RNN) is especially suitable to capture the temporal and spatial evolution of human moving and delay evolution after anomalies. Compared with a regression model, which restricts a constant relation between input and output, e.g., a polynomial relation, RNN presents higher flexibility on hidden relations. Besides, the configuration flexibility makes it suitable to integrate spatial and temporal dependency. However, previous studies [21] have shown that traditional RNNs fail to capture the long temporal dependency for the input sequence due to the vanishing gradient and exploding gradient problems. To address these drawbacks, Long Short-Term Memory (LSTM) is a special RNN architecture for sequence labeling and prediction tasks [21]. Therefore, we apply a time series learning LSTM model combined with spatial information and anomaly information to capture the delay time dynamics.

Fig. 45 illustrates the internal structure of LSTM cell, which consists of three gates, g_t is the input node at time t, which takes activation in the standard way from the input layer x_t and previous hidden layer h_{t-1}; i_t is the input gate, similar to g_t, which takes the input x_t and h_{t-1} with a sigmoid activation; s_t is a self-connected internal state with a fix unit weight, which is designed to solve the vanishing or exploding problem; f_t is the forget gate and used to flush the internal state; o_t is the output gate.

In our learning model, we construct a two-layer LSTM model including an encoding layer and decoding layer. The encoding layer takes two inputs $(d^{\mathcal{M}}_{status}(s, t_k), offset)$, the delay time and the offset of anomaly time, which is defined as $t_k - t_{\mathcal{E}}$ where $t_{\mathcal{E}}$ is the time when an anomaly happens.

Since we use sigmoid as activation function, the encoding layer captures the inter-dependency of delay and normalizes the output to a value between 0 and 1. In the decoding layer, we take the spatial information, anomaly information and the delay time inter-dependency as input. We apply a LSTM layer to predict the future delay time. The decoding layer is designed to capture the inter-dependency between delay time and external information, i.e., the anomaly type and the spatial information.

7.2 Evaluation Methodology

7.2.1 Learning Data. We build the learning model upon the dataset in our measurement study. The learning dataset includes four transportation systems, i.e., subway system, taxi system, bus system and private car system, which cover passengers' major modalities in a city. Each record in the dataset includes the following attributes:

location, *time slot*, *waiting time*,*riding time*, *transportation system*, and *anomaly*. We store the data in a relational database and build index on *location*, *time slot*, *transportation system* to construct travel time component tensors for efficient queries.

7.2.2 Cross Validation. We apply a 5-fold cross-validation, in each round, we use 1-fold as testing data. We train the models with different travel time components in multiple transportation systems. Specifically, when estimating delay time $d^{taxi}_{waiting}$, which is the delay waiting time in the taxi system, we train the model with historical $d^{taxi}_{waiting}$.

7.2.3 Metrics. We compare the predicted result $\hat{d}$ with the testing data $\bar{d}$ in terms of the delay time by Mean Absolute Percent Error (MAPE) defined in Equation 9.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|\hat{d}_i - \bar{d}_i|}{\bar{d}_i} \tag{9}$$

7.2.4 Baseline Approaches. We compare our model with three baseline methods.

- *ARIMA:* ARIMA is autoregressive integrate moving average model, which is a time series model to predict future points in the series. ARIMA model takes the delay time from previous n time slots as input and the delay time in the following time slot as the prediction target. To improve the prediction accuracy and reduce the noises introduced by anomaly types, we separate the data into groups by anomaly levels and categories and train a regression model on each data group. The external information, such as population and road networks, is not available in this prediction model [35].
- *MLE:* We implement the state-of-the-art model of travel time estimation [31], in which we fit a multinormal distribution in the anomaly area and surrounding regions. We set the delay time with different traveling paths as observations and maximize the likelihood of the estimated delay time based on the observations. We apply the Expectation-Maximization (EM) algorithm for the model convergence. The *MLE* model is implemented in different anomaly categories and levels.
- *MAC-:* We train three LSTM models: (i) a LSTM model without spatial information and anomaly information, (ii) a LSTM model with spatial information but without anomaly information, and (iii) a LSTM model with anomaly information but without spatial information. We integrate the three models and adopt the best performance as the output of *MAC-*.

7.2.5 Implementation. We implement our prediction model on travel time components inferred from the four transportation systems. Our model and baseline models are implemented with Keras and Tensorflow libraries. We train and evaluate our design on a server with 8 Nvidia K40C GPUs. We set the learning rate as 0.01. For each LSTM layer, we set the number of cells as 60 and initialize the LSTM parameter with random values between -0.01 to 0.01.

7.3 Evaluation Results

7.3.1 Performance. We predict the waiting delay and riding delay after an anomaly in the four transportation systems. Fig. 46 presents the performance of riding delay prediction after the anomaly for one-hour period. The *MLE* is a fitting model with historical delay time. However, different from travel time estimation, which shows regularity in the historical data, even though we categorize anomalies into different categories and levels and implement the model in a specific category and level, the delay time shows a high variance in anomalies. As a result, the delay time estimation depends on the time series features and time series models achieve better performance compared with the state-of-the-art travel time estimation model. The three time series models show similar performance trends in terms of time, where all models achieve the worse performance in the first

5 minutes. This is caused by two factors: (i) the uncertainty of the location, the level of the anomaly and the traffic condition, and (ii) the sparse after-anomaly observations as prediction features since *ARIMA* and *MAC-* take the previous delay time as input features for the prediction. After a half hour of an anomaly happened, the performances of the learning-based time series model become stable. Compared with the two baselines, our model achieves better performance in terms of the prediction accuracy and stability. The performance of the learning model on waiting delay prediction is given in Fig. 47, which shows similar results as the riding time estimation. However, the waiting time estimation error is larger compared with riding time.

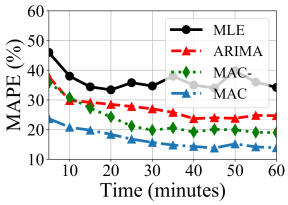

Fig. 46. Riding Delay

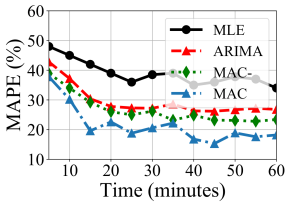

Fig. 47. Waiting Delay

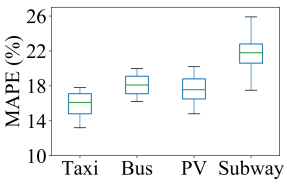

Fig. 48. Systems

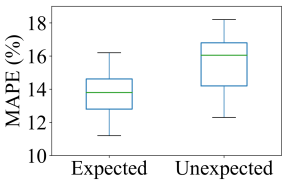

Fig. 49. Anomalies

7.3.2 Impact of Factors. We further investigate the impact of two factors on the prediction performance, i.e., the transportation systems and anomaly categories. Fig. 48 shows the performance distribution on the testing data. The box plot shows the minimum, first quarter, medium, third quarter and maximum of the distribution, which are used to present the average and the variance of the distribution. Even the subway system is the most stable in terms of travel time compared with above-ground systems in normal cases, among the four transportation systems, our algorithm has the lowest prediction error in the taxi system and the highest prediction error in the subway system. This is due to the anomalies in the subway system cause large-scale changes of traffic flow, especially the unexpected anomalies, which is difficult to be captured by limited observations. Comparing the three above-ground systems, we find that our algorithm achieves the best performance in the taxi system and it has the worst performance in the bus system. The reason is that the delay time of taxis is aggregated from vehicles, where much more taxis serve as travel time sensors, providing accurate observations as training features. Instead, since only riding time is predicted in private vehicles, the prediction error is the smallest in private vehicles among all above-ground systems.

8 DISCUSSIONS

Lessons learned: Based on the measurement results, we summarize a few lessons learned as follows. (i) Dividing travel time into fine-grained components helps us understand the impact of various factors on different parts of travel time, specifically urban anomalies. (ii) Compared with above-ground transportation systems, the anomalies have a larger impact on underground systems, i.e., subway systems, although underground systems are not affected by above-ground traffic condition. (iii) Unexpected anomalies have a larger impact on riding time and waiting time in all transportation systems compared with expected anomalies. (vi) While it is challenging to model and predict human mobility in terms of fine-grained travel time, spatial related contexts show strong correlations with human mobility and can improve the predictability significantly.

Generalization: We design, evaluate and implement the system in the Chinese city Shenzhen. Without access to the data in other cities, we cannot verify the effectiveness of our system in other cities. In particular, cities outside China have different policies for data releasing, which creates barriers to generalize our system in those cities with 4 transportation systems. However, since existing works have shown the accessibility to city-scale dataset in a single system such as London subway system [9], Beijing [26] and NYC [36] taxi system, Singapore

bus system [39], we believe the analysis method and prediction model can be generalized to other cities with similar spatio-temporal features. Moreover, our analysis about the impact of anomalies on different transportation modalities can be referred by city administrators and urban planners to better manage the city transportation. For the benefits of IMWUT community and peer researchers, we negotiate with the data provider and agree to release our sample dataset including one day of taxi GPS data, one day of subway smart card transaction records and one day of bus GPS records. Since private vehicle dataset contains personal information and is not a major design in our analysis, we will not release the private vehicle dataset.

Privacy Protections: While modeling the travel delay is important for individuals, we have to protect the privacy of participants involved. In this project, all data analyzed is anonymized by the collaborators, so data cannot be used to trace back to individual users explicitly. We only store and process data that is useful for the travel time modeling project, and exclude other information for the minimal exposure. All data are collected legally under the consent of the users.

Potential Societal Impacts: In the case study, we have shown that our measurement result could be utilized to predict the delay time when anomalies occur in a city. Our study can be applied to more potential applications with better societal impacts. (i) One application could be heterogeneous travel modes. Generally, people only take one mode of transportation towards their destinations if they do not have to transfer. Through understanding the travel delay in each area, people may be able to choose multiple transportation modes to their destinations depending on what kind of mode combinations achieve the minimum travel time. (ii) Another application could be the arrangement of public transportation. The travel delay of each transportation mode implies the balance of demand and supply in the area. Administrators can rearrange the existing transportation supply (e.g., increasing more buses in some lines) or guide the transportation setting (e.g., guiding the bus or subway line construction in the future).

9 CONCLUSION

In this work, we study the fine-grained travel time distribution in four transportation systems, which covers 10 million passengers in the Chinese city Shenzhen. In particular, we investigate the impact of anomalies on the travel time components in terms of expected anomalies and unexpected anomalies. Finally, we design a context-aware learning model to predict the fine-grained travel time under the impact of anomalies. Based on the above efforts, we provide a few valuable insights for fellow researchers to understand urban-scale human mobility with fine-grained travel behaviors. More importantly, our results have the potential to help the city government to manage urban traffic given expected anomalies and unexpected anomalies, which significantly improves urban efficiency and resilience.

ACKNOWLEDGMENTS

The authors would like to thank anonymous reviewers for their valuable comments and suggestions. This work is partially supported by NSF 1849238 and Rutgers Global Center.

REFERENCES

[1] Chao Chen, Daqing Zhang, Pablo Samuel Castro, Nan Li, Lin Sun, Shijian Li, and Zonghui Wang. 2013. iBOAT: Isolation-Based Online Anomalous Trajectory Detection. *IEEE Transactions on Intelligent Transportation Systems* 14, 2 (2013), 806–818. https://doi.org/10.1109/tits.2013.2238531

[2] Zhihan Fang, Fan Zhang, Ling Yin, and Desheng Zhang. 2018. MultiCell: Urban Population Modeling Based on Multiple Cellphone Networks. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 106 (Sept. 2018), 25 pages. https://doi.org/10.1145/3264916

[3] Piero Lovisolo Dario Parata Francesco Calabrese, Massimo Colonna and Carlo Ratti. 2011. Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS* 12 (March 2011).

[4] Mengdan Gao, Tongyu Zhu, Xuejin Wan, and Qi Wang. 2013. Analysis of travel time patterns in urban using taxi gps data. In *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social*

Computing. IEEE, 512–517.
[5] Qi Guan-De, Pan Yao, Li Shi-Jian, and Pan Gang. 2013. Predicting Passengers' Waiting Time by Mining Taxi Traces. (2013).
[6] Mordechai Haklay and Patrick Weber. 2008. Openstreetmap: User-generated street maps. *Ieee Pervas Comput* 7, 4 (2008), 12–18.
[7] Wonjae Jang. 2010. Travel time and transfer analysis using transit smart card data. *Transportation Research Record: Journal of the Transportation Research Board* 2144 (2010), 142–149.
[8] Yanjie Ji, Liangpeng Gao, Yingling Fan, Chu Zhang, and Ruochen Zhang. 2017. Waiting time perceptions at bus and metro stations in Nanjing, China: the importance of station amenities, trip contexts, and passenger characteristics. *Transportation Letters* (2017), 1–7.
[9] Neal Lathia and Licia Capra. 2011. How smart is your smartcard?: measuring travel behaviours, perceptions, and incentives. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 291–300.
[10] Neal Lathia and Licia Capra. 2011. Mining mobility data to minimise travellers' spending on public transport. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1181–1189.
[11] Haengju Lee, Desheng Zhang, Tian He, and Sang Hyuk Son. 2017. MetroTime: Travel Time Decomposition under Stochastic Time Table for Metro Networks. In *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*. 1–8.
[12] Junghoon Lee. 2008. Analysis on the Waiting Time of Empty Taxis for the Taxi Telematics System. In *2008 Third International Conference on Convergence and Hybrid Information Technology*, Vol. 1. 66–69.
[13] Bin Li, Daqing Zhang, Lin Sun, Chao Chen, Shijian Li, Guande Qi, and Qiang Yang. 2011. Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset. In *PerCom Workshops*.
[14] Miao Lin and Wen-Jing Hsu. 2014. Mining GPS data for mobility patterns: A survey. *Pervasive and Mobile Computing* 12 (2014), 1–16.
[15] Zhidan Liu, Zhenjiang Li, Mo Li, Wei Xing, and Dongming Lu. 2016. Mining road network correlation for traffic estimation via compressive sensing. *IEEE Transactions on Intelligent Transportation Systems* 17, 7 (2016), 1880–1893.
[16] Eric Hsueh-Chan Lu, Chih-Yuan Lin, and Vincent S Tseng. 2011. Trip-mine: An efficient trip planning approach with travel time constraints. In *Mobile Data Management (MDM), 2011 12th IEEE International Conference on*, Vol. 1. IEEE, 152–161.
[17] Google Map. 2018. Shenzhen Guangdong, China. (2018). https://www.google.ca/maps/place/Shenzhen,+Guangdong,+China/@22.5554167,113.9137924,11z/data=!3m1!4b1!4m5!3m4!1s0x3403f408d0e15291:0xfdee550db79280c9!8m2!3d22.543096!4d114.057865
[18] Stephan Wagner Gerd Kortuem Marcus Handte, Stefan Foell and Pedro José Marrón. 2016. An Internet-of-Things Enabled Connected Navigation System for Urban Bus Riders. *IEEE Internet of Things Journal* 3, 5 (October 2016).
[19] Population Reference Bureau. 2013. World Population Data Sheet. *http://www.prb.org/pdf13/2013-population-data-sheet-eng.pdf* (2013).
[20] Zhou Qin, Zhihan Fang, Yunhuai Liu, Chang Tan, Wei Chang, and Desheng Zhang. 2018. EXIMIUS: A measurement framework for explicit and implicit urban traffic sensing. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. ACM, 1–14.
[21] Xuan Song, Hiroshi Kanasugi, and Ryosuke Shibasaki. 2016. DeepTransport: Prediction and Simulation of Human Mobility and Transportation Mode at a Citywide Level.. In *IJCAI*, Vol. 16. 2618–2624.
[22] 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.
[23] Arvind Thiagarajan, James Biagioni, Tomas Gerlich, and Jakob Eriksson. 2010. Cooperative transit tracking using smart-phones. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. 85–98.
[24] Alejandro Tirachini. 2013. Estimation of travel time and the benefits of upgrading the fare payment technology in urban bus services. *Transportation Research Part C: Emerging Technologies* 30 (2013), 239–256.
[25] Huandong Wang, Fengli Xu, Yong Li, Pengyu Zhang, and Depeng Jin. 2015. Understanding mobile traffic patterns of large scale cellular towers in urban environment. In *Proceedings of the 2015 Internet Measurement Conference*. ACM, 225–238.
[26] Yilun Wang, Yu Zheng, and Yexiang Xue. 2014. Travel time estimation of a path using sparse trajectories. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 25–34.
[27] Xiaoyang Xie, Yu Yang, Zhihan Fang, Guang Wang, Fan Zhang, Fan Zhang, Yunhuai Liu, and Desheng Zhang. 2018. coSense: Collaborative Urban-Scale Vehicle Sensing Based on Heterogeneous Fleets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 196.
[28] Y. Zheng Y. Wang and Y. Xue. 2014. Travel Time Estimation of a Path using Sparse Trajectories. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2014).
[29] Yu Yang, Xiaoyang Xie, Zhihan Fang, Fang Zhang, Yang Wang, and Desheng Zhang. 2018. VeMo: Enabling Transparent Vehicular Mobility Modeling at Individual Levels with Full Penetration. *arXiv preprint arXiv:1812.02780* (2018).
[30] Yu Yang, Fan Zhang, and Desheng Zhang. 2018. SharedEdge: GPS-Free Fine-Grained Travel Time Estimation in State-Level Highway Systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 48.
[31] Yu Yang, Fan Zhang, and Desheng Zhang. 2018. SharedEdge: GPS-Free Fine-Grained Travel Time Estimation in State-Level Highway Systems. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 48 (March 2018), 26 pages. https://doi.org/10.1145/3191780
[32] Desheng Zhang, Jun Huang, Ye Li, Fan Zhang, Chengzhong Xu, and Tian He. 2014. Exploring human mobility with multi-source data at extremely large metropolitan scales. In *Proceedings of the 20th annual international conference on Mobile computing and networking*.

201–212.
[33] Desheng Zhang, Ye Li, Fan Zhang, Mingming Lu, Yunhuai Liu, and Tian He. 2013. coRide: Carpool Service with a Win-win Fare Model for Large-scale Taxicab Networks. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys '13)*. ACM, New York, NY, USA, Article 9, 14 pages. https://doi.org/10.1145/2517351.2517361
[34] Desheng Zhang, Juanjuan Zhao, Fan Zhang, and Tian He. 2015. UrbanCPS: A Cyber-physical System Based on Multi-source Big Infrastructure Data for Heterogeneous Model Integration. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems (ICCPS '15)*. ACM, New York, NY, USA, 238–247. https://doi.org/10.1145/2735960.2735985
[35] Huichu Zhang, Yu Zheng, and Yong Yu. 2018. Detecting Urban Anomalies Using Multiple Spatio-Temporal Data Sources. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 54.
[36] Jianting Zhang. 2012. Smarter outlier detection and deeper understanding of large-scale taxi trip records: a case study of NYC. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*. 157–162.
[37] Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, and Wei-Ying Ma. 2008. Understanding mobility based on GPS data. In *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 312–321.
[38] Yu Zheng, Huichu Zhang, and Yong Yu. 2015. Detecting collective anomalies from multiple spatio-temporal datasets across different domains. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2.
[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*. ACM, 379–392.
[40] Brian D. Ziebart, Andrew L. Maas, Anind K. Dey, and J. Andrew Bagnell. 2008. Navigate like a cabbie: probabilistic reasoning from observed context-aware behavior. In *Proceedings of the 10th international conference on Ubiquitous computing (UbiComp '08)*.

Received November 2018; revised February 2019; accepted April 2019