

iTV: Inferring Traffic Violation-Prone Locations with Vehicle Trajectories and Road Environment Data

Zhihan Jiang, Longbiao Chen, *Member, IEEE*, Binbin Zhou, Jinchun Huang, Tianqi Xie, Xiaoliang Fan, *Member, IEEE*, Cheng Wang, *Member, IEEE*

Abstract—Traffic violations like illegal parking, illegal turning, and speeding have become one of the greatest challenges in urban transportation systems, bringing potential risks of traffic congestions, vehicle accidents, and parking difficulties. To maximize the utility and effectiveness of the traffic enforcement strategies aiming at reducing traffic violations, it is essential for urban authorities to infer the traffic violation-prone locations in the city. Therefore, we propose a low-cost, comprehensive, and dynamic framework to infer traffic violation-prone locations in cities based on the large-scale vehicle trajectory data and road environment data. Firstly, we normalize the trajectory data by map matching algorithms and extract key driving behaviors, i.e., turning behaviors, parking behaviors, and speeds of vehicles. Secondly, we restore spatiotemporal contexts of driving behaviors to get corresponding traffic restrictions such as no parking, no turning, and speed restrictions. After matching the traffic restrictions with driving behaviors, we get the traffic violation distribution. Finally, we extract the spatiotemporal patterns of traffic violations, and build a visualization system to showcase the inferred traffic violation-prone locations. To evaluate the effectiveness of the proposed method, we conduct extensive studies on large-scale, real-world vehicle GPS trajectories collected from two Chinese cities, respectively. Evaluation results confirm that the proposed framework infers traffic violation-prone locations effectively and efficiently, providing comprehensive decision supports for traffic enforcement strategies.

Index Terms—traffic violation, vehicle trajectory data, traffic sign detection, map matching, crowdsensing.

I. INTRODUCTION

TRAFFIC violations, such as speeding and illegal parking, have become one of the greatest challenges in urban transportation systems, bringing potential risks of traffic congestions, vehicle accidents, and parking difficulties, etc. [1], [2], [3]. For example, in 2018, New York City witnessed 54,469 traffic violations and 44,508 traffic injuries across the city [4]. To reduce traffic violations, urban authorities have implemented various traffic enforcement strategies, such as deploying field enforcement officers in rush hours and installing

Manuscript received March 14, 2020; revised June 11, 2020; accepted July 21, 2020. Date of publication ...; date of current version July 27, 2020. (Corresponding author: Longbiao Chen)

Z. Jiang, L. Chen, J. Huang, T. Xie, X. Fan and C. Wang are with Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China (e-mail: zhihanjiang@stu.xmu.edu.cn, longbiaochen@xmu.edu.cn).

B. Zhou is with Zhejiang University, Hangzhou 310000, China (e-mail: bbzhou@zju.edu.cn).

Digital Object Identifier 10.1109/JSYST.2020.3012743

Copyright: 0000-0000/00\$00.00 © 2020 IEEE

automated monitoring cameras in road intersections [5]. Given the expensive human resource allocation and infrastructure investment, it is essential for urban authorities to identify the traffic violation-prone locations so as to deploy officers and install cameras under limited labor and non-labor resources.

However, traditional strategies for traffic violation-prone location inference are highly dependent on historical traffic violation records and human experience, which are labor intensive, time consuming, and unable to adapt to rapidly developing cities. Therefore, a low-cost, comprehensive, and dynamic method is in great demand. Fortunately, with the popularization of GPS devices and map services like street view service, we can get crowd-sensed and large-scale vehicle trajectory data in cities, real panoramic street view pictures on roads, and related traffic restrictions such as speed restrictions. These rich trajectory data and road environment data provide us an unprecedented opportunity to explore traffic violation-prone locations.

In this work, we propose a low-cost, comprehensive and dynamic framework for inferring the traffic violation-prone locations in cities based on the crowd-sensed, large-scale vehicle trajectory data and road environment data fusion, so that we can provide some insights for the traffic management department about traffic violation-prone locations to help optimize the utility and effectiveness of the traffic enforcement strategies.

Firstly, we normalize the trajectory data by mapping the vehicle trajectories onto the road network and get the driving behaviors. Secondly, we model driver perspectives to match driving behaviors to corresponding road segments and get the spatiotemporal contexts of driving behaviors. Using the spatiotemporal context, we detect traffic signs to identify no-turning road intersections and no-parking road segments. We can also get speed restrictions on roads from real-time navigation service providers. After matching the traffic restriction information with driving behaviors, we extract three types of traffic violations, i.e., illegal turning, illegal parking and speeding, and extract the spatiotemporal patterns of traffic violations to infer the traffic violation-prone locations. Finally, we build a traffic violation-prone locations inference system and evaluate the proposed method using large-scale, real-world datasets from two cities in China, Chengdu and Xiamen.

In designing the framework, there are several research issues to be addressed:

- 1) It is non-trivial to extract turning behaviors from

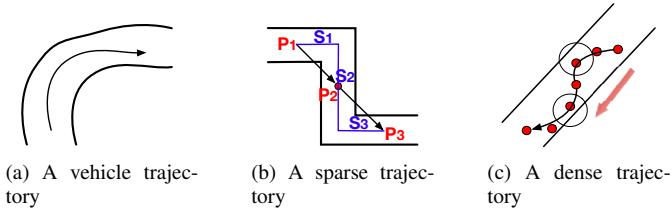


Fig. 1: Some misunderstanding examples of turning behaviors.

vehicle GPS trace data. When only trajectory data is used, it is easy to misunderstand some driving behaviors. For example, as shown in Fig. 1a, if we observe the trajectory separately, we may consider that it is a turning behavior in an intersection. However, if we locate the trajectory into the corresponding road network, we can find that it is not a turning behavior since it is caused by the curvy road rather than the driver's decision in the intersection. Moreover, some sparse GPS points will also mislead us. As shown in Fig. 1b, the trajectory $P_1 \rightarrow P_2 \rightarrow P_3$ implies us a straight line, while it should be $S_1 \rightarrow S_2 \rightarrow S_3$ actually when combined with the road network. Here, two turning behaviors can be extracted. Besides, if the GPS points are dense and the driver prefers changing lanes while driving, there will be several curves on the trajectory, as shown in the circled parts in Fig. 1c. Then those curves will be mistaken for turning behaviors. Therefore, to extract turning behaviors accurately, we should take both GPS trajectories and road networks into consideration.

- 2) **It is difficult to restore the spatiotemporal contexts of driving behaviors.** To identify whether a driving behavior is illegal or not, we need to restore the spatiotemporal context of driving behavior. For example, a driver makes an impermissible left turn in the intersection with a no-left-turn sign located in the driver's previous perspective; thus, this behavior can be identified as a traffic violation. However, with only vehicle GPS trajectories, it is difficult to get traffic restriction information for road segments where the trajectories are. Fortunately, the development of various map services enables us to get panoramic pictures of the surrounding environment of almost all streets in the city and real-time navigation service can offer us speed restriction information while driving. Those services enable us to restore spatiotemporal contexts of driving behaviors.

In summary, the main contributions of this paper include:

- 1) To the best of our knowledge, this is the first work using crowd-sensed and large-scale vehicle trajectory and road environment data to infer traffic violation-prone locations in cities. Such a low-cost, dynamic, comprehensive method can help urban authorities maximize the utility and effectiveness of the traffic enforcement strategies.
- 2) We proposed a data-driven method to identify traffic violation-prone locations. Firstly, we normalize the GPS trajectories with map matching methods and get the driving behavior distribution. Secondly, we model driver perspectives based on regression to augment the spatiotemporal contexts of driving behaviors. Thirdly, we extract three types of traffic restrictions from those contexts, i.e., no-parking, no-turning, and speed restric-

tions. After matching those restrictions with driving behaviors, we get the traffic violation distribution and extract the spatiotemporal features of traffic violations. Finally, we build the traffic violation-prone location inference system.

- 3) We evaluate the proposed method using large-scale, real-world datasets from two Chinese cities, i.e., Chengdu and Xiamen, including vehicle GPS trajectories, street view pictures, and speed restrictions. The experimental results are visualized in a traffic violation-prone location inference system.

II. RELATED WORK

A. GPS Trajectory Mining

In the literature, there have been many studies on mining various kinds of GPS trajectories for different application scenarios [6]. In a study on human mobility, Zheng et al. [7] proposed a framework to mine interesting locations and travel sequences from human GPS trajectories, while Alvares et al. [8] proposed a model to enrich human GPS trajectories with semantic geographical meanings. On taxi operation study, Zhang et al. [9] detected anomalous passenger delivery trips from taxi GPS traces, Chen et al. [10] explored citywide night bus planning issues leveraging taxi GPS traces, and Zhang et al. [11] identified taxi refueling behaviors from GPS trajectories to estimate citywide petrol consumption and analyze gas station efficiency. Chen et al. [12] proposed a framework for container port performance measurement and comparison using ship GPS traces. As for driving violation identification, He et al. [13] proposed a framework to detect illegal parking events using shared bikes' trajectories. Lee et al. [14] proposed a methodology for detecting illegal parking events in real-time by applying a novel picture projection that reduces the dimensionality of the data. Cerber et al. [15] proposed a disclosed method of traffic control, which can detect and identify the vehicle speeding. In this work, the GPS trajectory data is one of the major resources, on which we can infer the traffic violation-prone locations.

B. Traffic Sign Detection

Traffic sign detection has been widely studied in the field of computer vision. Escalera et al. [16] used color thresholding and shape analysis to detect signs in pictures and classified signs with a neural network. Bahlmann et al. [17] detected signs using a set of Haar wavelet features obtained from AdaBoost training and classified signs using Bayesian generative modeling. Mogelmose et al. [18] provided a survey of the traffic sign detection literature, detailing detection systems for traffic sign detection for driver assistance. Houben et al. [19] introduced a real-world benchmark data set from Germany for traffic sign detection and compared several detection approaches. In recent years, deep learning methods have shown superior performance for many tasks, such as object detection. Zhu et al. [20] proposed a framework of traffic sign detection and detection based on proposals by the guidance of a fully convolutional network. Zhu et al. [21] provided 100000 pictures containing 30000 traffic-sign instances and

demonstrated how a robust end-to-end convolutional neural network (CNN) could simultaneously detect and classify traffic signs. Luo et al. [22] proposed a data-driven system to detect traffic signs, which include both symbol-based and text-based signs, in video sequences captured by a camera mounted on a car. Huang et al. [23] introduced GAN into the Faster-RCNN framework and improved the performance of small object detection compared to Faster-RCNN. Yolov3 [24] is a well-known state-of-art object detection system, which has achieved great performance both in accuracy and time. It has made great progress in small object detection and has strong generalization ability. In this work, we choose YOLOv3 as the backbone network to train the traffic sign detection model with transfer learning [25]. We first use the Chinese traffic-sign benchmark provided by [21] to pre-train a traffic sign detection model. Then we fine-tune the model based on a small dataset with less categories of traffic signs collected from the targeted city. In this way, the final model can be adapted to the task of interested categories of traffic signs and cities, and it can be transferred to new cities easier, since it only requires a small dataset to fine-tune the pre-trained model.

C. Map Matching Methods

Map matching is the procedure for mapping vehicle trajectories onto the correct road segments, which plays an important role in-vehicle navigation systems, routing engines, etc. [26]. Ren et al. [27] introduced the GPS-based wheelchair navigation based on a novel map matching algorithm. Bernstein et al. [28] introduced the personal navigation assistants based on the map matching method. He et al. [13] used the map matching method to detect illegal parking events. White et al. introduced some map matching algorithms for personal navigation assistant [29]. Chen et al. [30] proposed an online map-matching-based trajectory compression framework running under the mobile environment. Vaughn et al. [31] described the GPS-map speed matching system for controlling the speed of the vehicle. The nearest matching method is the most basic map matching method, in which the latitude/longitude pairs are assigned to the nearest points or curves on the road network. However, this method is unable to take the contexts of GPS point sequences into account. Based on the nearest matching method, some environment factors, such as orientation information, connectivity constraints, and topologically feasible path through the road network, are incorporated to improve the performance of map matching [29], [32]. However, these purely geometric methods are sensitive to measurement noise and sampling rate, while the Hidden Markov Model (HMM) addresses this issue by explicitly modeling the connectivity of the roads and considering many different path hypotheses simultaneously [26], [33], [34]. In this work, we normalize the trajectory data with a map matching method so as to extract driving behaviors, and since the trajectory data are noisy and nonuniform in real-world deployment, we choose to use the map matching algorithm based on HMM.

III. PRELIMINARY AND FRAMEWORK OVERVIEW

In this section, we introduce different kinds of traffic signs, define several terms used in this paper, and present the



Fig. 2: Four most frequently used traffic sign types. (a) No-left-turn. (b) No-right-turn. (c) No-u-turn. (d) No-parking.

overview of the proposed framework.

A. Preliminary

Road Network: A road network can be defined as a graph $G = (I, R)$, where $I = \{i_1, i_2, \dots, i_n\}$ is a set of road intersections, and $R = \{r_1, r_2, \dots, r_m\}$ is a set of road segments.

GPS Point Trajectory: A GPS point can be denoted as 4-tuples, $p = (id, t, lat, lng)\}$, where id, t, lat, lng are the unique trajectory ID, time stamp, latitude, and longitude from GPS transmitters, respectively. A GPS point trajectory $traj_p$ can be defined as a time-ordered sequence $P_s = \{p_1 \rightarrow p_2 \rightarrow \dots \rightarrow p_n\}$, where n is the number of GPS points in the trajectory.

Road Segment Trajectory: A road segment trajectory $traj_g$ can be denoted by 2-tuples, $traj_g = (id, R_s)$, where id is the unique trajectory ID, and R_s is defined as a time-ordered sequence of road segments, $R_s = \{r_1 \rightarrow r_2 \rightarrow \dots \rightarrow r_m\}$, where $r_i \in R, 1 \leq i \leq m$ and m is the number of road segments in the trajectory.

Turning Behavior: The turning behavior can be defined as 8-tuples, $tn_i = (type, id, lat, lng, t, bb, ba, conf)$, where $id, lat, lng, t, bb, ba, conf$ are the unique trajectory ID, latitude, longitude, time of the turning behavior, bb and ba are the direction of the vehicle before and after the turning behavior, respectively, which are the clockwise angles from the North, $conf$ is the confidence value of the turning behavior and $type$ is the type of the turning behavior, $type \in \{left - turn, right - turn, u - turn\}$, corresponding to the three traffic signs shown in Fig. 2. A set of turning behaviors is denoted by $TN = \{tn_1, tn_2, \dots, tn_m\}$, where m is the number of turning behaviors.

Parking Behavior: The parking behavior can be defined as 5-tuples, $pk_i = (id, lat, lng, st, et)$, where id, lat, lng, st, et are the unique trajectory ID, longitude, latitude, start time and end time of the parking behavior, which is consistent with the no-parking sign shown in Fig. 2 (d). A set of parking behaviors is denoted by $PK = \{pk_1, pk_2, \dots, pk_m\}$, where m is the number of parking behaviors.

B. System Framework Overview

As shown in Fig. 3, the framework consists of three phases, i.e., driving behavior extraction, driving behavior context augmentation and traffic violation-prone location inference system. We briefly elaborate on the whole process as follows.

In the driving behavior extraction phase, we normalize the trajectory data based on the road network using map matching methods. Then we extract parking behaviors from the normalized data by static points and extract turning behaviors from the road segment trajectory. In the driving behavior context augmentation phase, we model driver perspectives

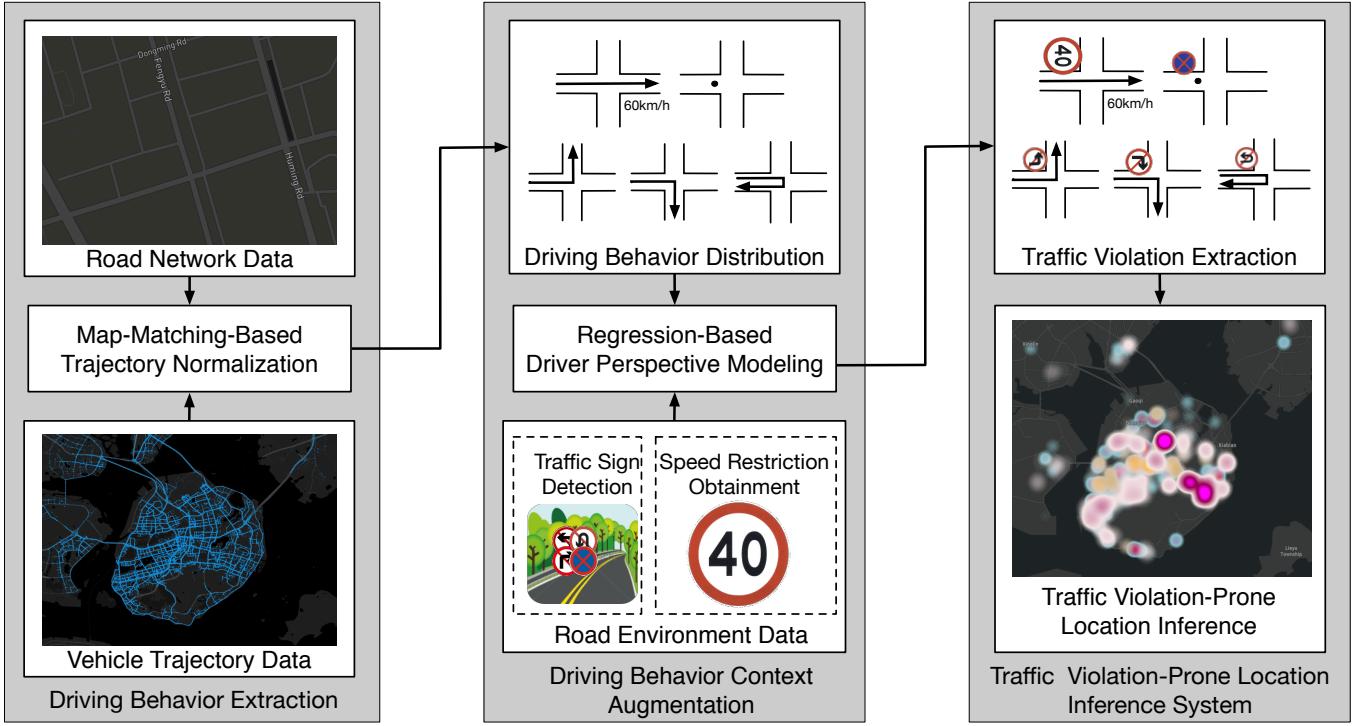


Fig. 3: Framework Overview.

based on regression and retrieve corresponding street view pictures and other traffic restriction information. Then we train a traffic sign detection model to detect traffic signs in street view pictures and obtain speed restrictions from real-time navigation service providers. In the traffic violation-prone location inference system phase, we first match driving behaviors with those restrictions to get the traffic violation distribution. Then we extract the spatiotemporal patterns of traffic violations to infer the traffic violation-prone locations and build a traffic violation-prone location inference system.

IV. DRIVING BEHAVIOR EXTRACTION

In this section, our goal is to extract average velocities and driving behaviors, i.e., parking, left turn, right turn, and u-turn from crowd-sensed and large-scale vehicle GPS trajectories. However, the GPS readings are usually noisy and nonuniform due to many factors such as poor conditions of GPS devices and different sampling rates [35], and it is non-trivial to extract driving behaviors directly from the trajectories without the information about road network. To address these challenges, we first employ a map matching method to normalize trajectory data into road segment trajectories, and then extract the driving behaviors, i.e., parking, left turn, right turn, and u-turn from the normalized trajectories.

A. Map-Matching-Based Trajectory Normalization

The purely geometric map matching methods are sensitive to measurement noise and sampling rate, while the trajectory data are usually noisy and nonuniform in real-world deployment. Therefore, in this paper, a HMM-based map matching algorithm which addresses this issue is used to find the most

likely road route represented by a timestamped sequence of latitude/longitude pairs [26].

We first preprocess the raw vehicle trajectories by removing duplicate and abnormal points and reconstructing trajectories. In this algorithm, the states of the HMM are the individual road segments, denoted as r_i , $i = 1, \dots, m$, and the state measurements are the noisy vehicle location measurements. The goal is to match each latitude/longitude location measurement z_t with the proper segment r_i .

The emission probability for each road segment r_i and each location measurement z_t is $p(z_t|r_i)$, which gives the likelihood that z_t would be observed if the vehicle were actually on road segment r_i , and the closest point on the road segment is denoted as $x_{t,i}$, and direct distance (the great circle direct distance on the surface of the earth between two GPS points) between the measured point and the candidate match is $\|z_t - x_{t,i}\|_d$. The intuition is that road segments farther from the measurement are less likely to have produced the measurement. The GPS noise are modeled as zero-mean Gaussian [36], i.e.,

$$p(z_t|r_i) = \frac{1}{\sqrt{2\pi\sigma_z^2}} e^{-0.5(\frac{\|z_t - x_{t,i}\|_d}{\sigma_z})^2} \quad (1)$$

where σ_z is the standard deviation of GPS measurements, and the points within $2\sigma_z$ of the previously included point are removed. σ_z is estimated by the median absolute deviation [26], i.e.,

$$\sigma_z = 1.4826 \times \text{median}_t(\|z_t - x_{t,i}\|_d) \quad (2)$$

The initial state probabilities are started at the first measurement, $\pi_i = p(z_1|r_i)$, $i = 1, \dots, m$. Also, we do not consider matching to road segments that are quite distant from the measurement. The measurement probability from a road

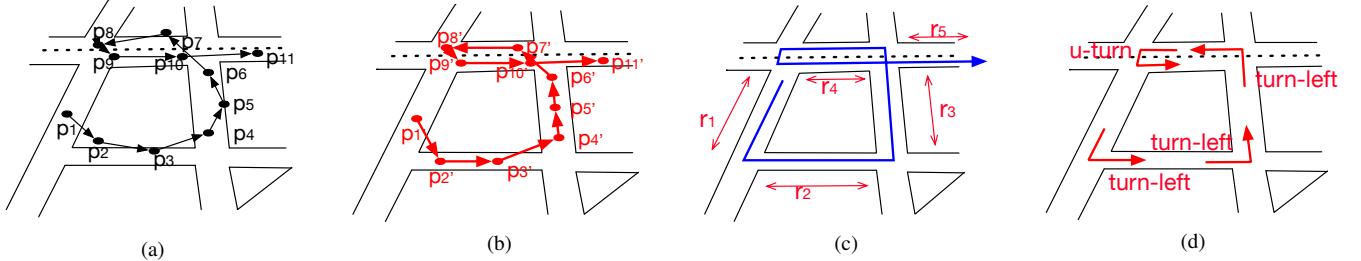


Fig. 4: An illustration of map-matching-based trajectory normalization. (a) A GPS point trajectory. (b) A normalized GPS point trajectory. (c) A normalized road segment trajectory. (d) A turning behavior trajectory.

segment that is more than 200 meters away from z_t is set to zero, which helps reduce the number of candidate matches.

The transition probabilities give the probability of a vehicle moving between the candidate road matches at t and $t+1$. Specifically, for the next measurement z_{t+1} and candidate road segment r_j , the corresponding point is $x_{t+1,j}$. We compute the direct distance between z_t and z_{t+1} , $\|z_t - z_{t+1}\|_d$ and the route distance (the shortest length of road segments between two GPS points on the road) between x_t and x_{t+1} , $\|x_{t,i} - x_{t+1,j}\|_r$. With the intuition that these two distances will be about the same for correct matches, and the absolute values of these two distances from the correct matches fits well to an exponential probability distribution given by [26]

$$p(d_t) = \frac{1}{\beta} e^{-\frac{d_t}{\beta}} \quad (3)$$

where $d_t = \|x_{t,i} - x_{t+1,j}\|_r - \|z_t - z_{t+1}\|_d$, and β is estimated by a robust estimator suggested by Gather and Schultze [37], i.e.,

$$\beta = \frac{1}{\ln(2)} \text{median}_t(|(\|z_t - z_{t+1}\|_d - \|x_{t,i} - x_{t+1,j}\|_r)|) \quad (4)$$

We set $p(d_t)$ to zero when d_t is greater than 2000 meters and terminate the search for a route. And if a calculated route would require the vehicle to exceed a speed of 180 kilometers per hour, we set its probability to zero. With the measurement probabilities and transition probabilities, the Viterbi algorithm [38] is used to compute the best path.

After map matching, the GPS point trajectory $traj_p = (id, P_s)$ is normalized into road segment trajectory $traj_g = (id, R_s)$. For example, as shown in Fig. 4a, there is a GPS point trajectory $traj_p = (i, P_s)$, where $P_s = \{p_1 \rightarrow p_2 \rightarrow p_3 \rightarrow p_4 \rightarrow p_5 \rightarrow p_6 \rightarrow p_7 \rightarrow p_8 \rightarrow p_9\}$. After map matching, we get the corresponding locations of each GPS points on the road segment, as shown in Fig. 4b. The mapped point p'_1 is on road segment r_1 , p'_2 and p'_3 are on road segment r_2 , p'_4 to p'_6 are on road segment r_3 , p'_7 to p'_{10} are on r_4 , p'_1 is on r_5 . Thus we can get the corresponding road segment trajectory $traj_g = (i, R_s)$, where $R_s = \{r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow r_4 \rightarrow r_4 \rightarrow r_5\}$, as shown in Fig. 4c.

B. Driving Behavior Extraction

Based on the normalized trajectories, we further extract vehicles' average velocities on roads, the turning behaviors in road intersections, and the parking behaviors on road segments from the normalized trajectories.

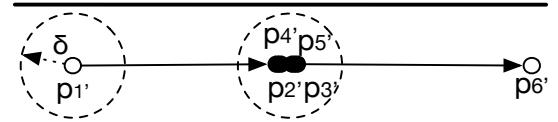


Fig. 5: An illustration of extracting the parking behavior.

1) *Turning Behavior Extraction*: From the normalized road segment trajectories, we extract the turning behaviors. bb and ba are the directions of the vehicle before and after the turning behavior, which are the clockwise angles from the North. If $ba - bb = 0^\circ$, it is a straight behavior, which is not considered in this paper; if $0^\circ < ba - bb < 160^\circ$, it is a right-turn behavior; if $160^\circ \leq ba - bb \leq 200^\circ$, it is a u-turn behavior; if $200^\circ < ba - bb < 360^\circ$, it is a left-turn behavior. As shown in Fig. 4, from r_1 to r_2 , there is a left-turn behavior; from r_2 to r_3 , there is a left-turn behavior; from r_3 to r_4 , there is a left-turn behavior; from r_4 to r_5 , there is a u-turn behavior; from r_4 to r_5 , it is a straight behavior. Therefore we can get a turning behavior trajectory, $\{\text{left-turn} \rightarrow \text{left-turn} \rightarrow \text{left-turn} \rightarrow \text{u-turn}\}$, as shown in Fig. 4 (d).

2) *Parking Behaviors Extraction*: We extract the parking behaviors in the normalized GPS point trajectories with a sliding-window-based method [12]. More specifically, for a normalized GPS point trajectory $traj_p = (i, P_s)$, where $P_s = \{p'_1 \rightarrow p'_2 \rightarrow \dots \rightarrow p'_n\}$, we extract every parking sequence $p'_m \rightarrow p'_{m+1} \rightarrow \dots \rightarrow p'_{m+k}$ ($1 \leq m < n, 1 \leq k \leq n - m$) in which the average speed between the first point and any other points is less than a small threshold δ [12], i.e.,

$$\forall m \leq i < m + k, \frac{dist(p'_i, p'_{i+1})}{\Delta t} < \delta \quad (5)$$

where $dist(p'_i, p'_{i+1})$ is the route distance between p'_i and p'_{i+1} and $\Delta t = |p'_i.t - p'_{i+1}.t|$ is the time difference of p'_i and p'_{i+1} . We use a sliding-window with adaptive size along the trajectory to find such parking behaviors. In particular, we dynamically extend the window size by adding new points until the newly-formed sequence violates requirement 5. For example, as shown in Fig. 5, for the trajectory $p'_1 \rightarrow p'_2 \rightarrow p'_3 \rightarrow p'_4 \rightarrow p'_5 \rightarrow p'_6$, we start by creating a window consisting of the first two points (p'_1, p'_2 in this case), and check whether the average speed between p'_1 and p'_2 is less than δ . Since $\frac{dist(p'_1, p'_2)}{|p'_1.t - p'_2.t|} > \delta$, we discard this window, and slide the window to start over from the end point(p'_2), and create a new window (p'_2, p'_3). We keep this window for $\frac{dist(p'_2, p'_3)}{|p'_2.t - p'_3.t|} < \delta$ and repeat the procedure for the next adjacent points until the speed

constraint is violated. Finally, we obtain a sequence containing a set of consecutive points $p'_2 \rightarrow p'_3 \rightarrow p'_4 \rightarrow p'_5$.

We map each parking sequence $p'_m \rightarrow p'_{m+1} \rightarrow \dots \rightarrow p'_{m+k}$ extracted from the GPS trajectories to a parking behavior $pk = (id, lat, lng, st, et)$. The id of pk is the id of the trajectory where the parking behavior is extracted, lat is the average latitude of the points in the parking sequence, $lat = \frac{\sum_{i=m}^{m+k} p'_i.latitude}{k+1}$, lng is the average longitude of the points in the parking sequence, $lng = \frac{\sum_{i=m}^{m+k} p'_i.longitude}{k+1}$, st is start time of the parking behavior, $st = p'_m.t$, et is the end time of the parking behavior, $et = p'_{m+k}.t$.

3) *Average Velocity Extraction*: After trajectory normalization, we can get sequences of road intersections, from which we extract the subsequences. The intersections from the same subsequence are of the same road. For each subsequence, we calculate the average speed through dividing the route distance between the head intersection and the rear intersection by the difference of the time. More specifically, for a road intersection sequence $p_1 \rightarrow p_2 \rightarrow p_3 \rightarrow \dots \rightarrow p_n$, where n is the length of the sequence. If p_i, p_{i+1} and p_{i+2} are of the same road, we extract the subsequences $p_i \rightarrow p_{i+1} \rightarrow p_{i+2}$, and the average speed $v = \frac{dist(p_i, p_{i+2})}{\Delta t}$, where $dist(p_i, p_{i+2})$ and Δt are the route distance and duration between p_i and p_{i+2} .

V. DRIVING BEHAVIOR CONTEXT AUGMENTATION

After extracting driving behaviors, we need to know the contexts of driving behaviors to identify whether a driving behavior is illegal or not. To this end, we model the driver perspectives based on regression and get the corresponding street view pictures to detect the no-left-turn, no-right-turn, no-u-turn, and no-parking signs.

A. Driver Perspective Modeling

Since vehicle trajectory data we use in this paper are extremely large-scale, the GPS points in the trajectories can almost cover the urban road network completely. Therefore, after getting the driving behavior database $DB = (PK, TN)$, we can detect almost all the road intersections in the city by extracting the distinct location of turning behaviors. The no-left-turn, no-right-turn, and no-u-turn traffic signs are erected near the road intersections to guide drivers. And the no-parking signs are often erected on the road intersections to inform drivers whether they can park on the following road segments.

For each intersection, there are several different kinds of turning behaviors from different directions. We define those turning behaviors at the same road intersection from the same direction as a *bunch* = (lat, lng, bb) , where lat, lng are the latitude and the longitude of the intersection, bb is the bearing before the behaviors. Thus for the turning behaviors from one *bunch*, their spatiotemporal contexts before the behaviors should be the same. As shown in Fig. 6, the behaviors drew in red are a *bunch*, denoted as *bunch_r* and those in blue are another *bunch*, denoted as *bunch_b*.

For each $bunch_i = (lat_i, lng_i, bb_i)$, we select the corresponding GPS point trajectories in *bunch_i*, and for each trajectory j , the corresponding turning behavior at this road

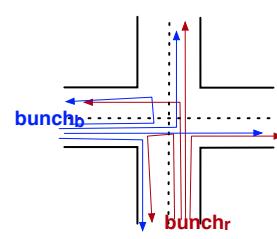


Fig. 6: An example of two *bunches* of trajectories.

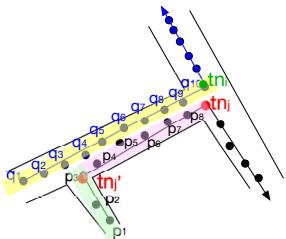


Fig. 7: An illustration of the driver perspective modeling.

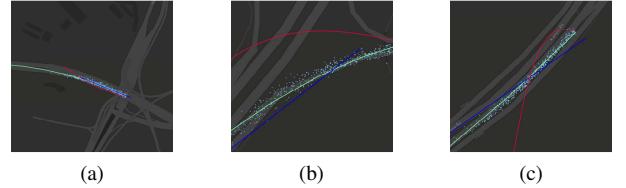


Fig. 8: The results of linear regression (the blue lines), spline regression (the red lines) and cubic polynomial regression (the green lines).

intersection is tn_j , and we select a segment of the trajectory $traj_{pj} = \{p_k \rightarrow p_{k+1} \rightarrow \dots \rightarrow p_{k+m}\}$ before tn_j , the time span of which is limited by a threshold θ . Moreover, in order to avoid the GPS points on the branch road, the start time of the trajectory segment should be no earlier than the former turning behavior tn'_j of tn_j , i.e.,

$$tn_j.t - p_k.t < \theta, p_{k+m} \leq tn_j.t, tn'_j.t \leq p_k.t \leq tn_j.t \quad (6)$$

For example, as shown in Fig. 7, suppose that there are two GPS trajectories in the *bunch* at this road intersection, according to 6, we select two trajectory segments, $traj_{p1} = \{q_1 \rightarrow q_2 \rightarrow q_3 \rightarrow q_4 \rightarrow q_5 \rightarrow q_6 \rightarrow q_7 \rightarrow q_8 \rightarrow q_9 \rightarrow q_{10}\}$ and $traj_{p2} = \{p_1 \rightarrow p_2 \rightarrow p_3 \rightarrow p_4 \rightarrow p_5 \rightarrow p_6 \rightarrow p_7 \rightarrow p_8\}$, where $tn_i.t - q_1.t < \theta, q_{10} \leq tn_i.t, tn_j.t - p_1.t < \theta$ and $p_8 \leq tn_j.t$. However, for $traj_{p2}$, there is a right-turn behavior tn'_j before tn_j , $p_1.t < tn'_j.t$ and $p_2.t < tn'_j.t$, so we drop points p_1 and p_2 . Finally, the trajectory segments selected are $q_1 \rightarrow q_2 \rightarrow q_3 \rightarrow q_4 \rightarrow q_5 \rightarrow q_6 \rightarrow q_7 \rightarrow q_8 \rightarrow q_9 \rightarrow q_{10}$ and $p_3 \rightarrow p_4 \rightarrow p_5 \rightarrow p_6 \rightarrow p_7 \rightarrow p_8$, respectively.

Since the trajectory data are extremely massive, we select a large number of trajectory segments that can describe contexts before turning behaviors at the road intersections. For each turning behaviors tn_i , we denote the GPS points in the corresponding trajectory segments as a set $P_i = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, where m is the number of the points, we model the shape of the lane segment by regression. We select five typical road segments and try three kinds of regression, linear regression, spline regression, and cubic polynomial regression. The result shows that cubic polynomial regression achieves the best performance, as shown in Fig. 8.

Therefore for $(x_i, y_i) (i = 1, \dots, m)$, we can get a curve $h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$ satisfied that [39], [40],

$$\theta_0, \theta_1, \theta_2, \theta_3 = \arg \min_{\theta_0, \theta_1, \theta_2, \theta_3} \frac{1}{2m} \sum_{i=1}^m (h(x_i) - y_i)^2 \quad (7)$$

The regression curve segment can be regarded as a virtual driver's route before the driving behavior, representing all drivers of the selected trajectories.



Fig. 9: An example of a sequence of street view pictures.



Fig. 10: The types of traffic signs selected from the Chinese traffic-sign benchmark.

B. Driving Behavior Context Augmentation

We pick several points as a point sequence on the regression curve segment uniformly and normalized them by map matching, regard the tangent of the curve at those points as the direction of the vehicle at these locations. Therefore we can get the corresponding street view picture sequence, through which we can restore what drivers have seen before those road intersections. For example, Fig. 9 shows a sequence of street view pictures. In addition, we also collect the speed restrictions of each road from real-time navigation service providers.

VI. TRAFFIC VIOLATION-PRONE LOCATION INFERENCE SYSTEM

In this section, our goal is to infer the traffic violation-prone locations in cities. However, since the intelligent information management system have not been implemented in many developing or developed cities, the traffic management departments usually do not have a comprehensive knowledge about the traffic restriction distribution in the whole city. Therefore, we obtain the traffic rule information through detecting traffic signs in street view pictures so as to identify traffic violations. Then we extract the spatiotemporal patterns of traffic violations to infer traffic violation-prone locations.

A. Traffic Sign Detection

From driver perspectives, we can find no-left-turn signs, no-right-turn signs, no-u-turn signs, and no-parking signs erected at the side of or above the roads to give instructions to drivers. After modeling driver perspectives, we then detect these four categories of traffic signs in the spatiotemporal context around the road intersections to help infer the illegal turning behaviors.

The state-of-art object detection system YOLOv3 [24] has achieved great performance both in accuracy and time. It has a few incremental improvements on YOLOv2 [41], such as using independent logistic classifiers instead of softmax, adding shortcut connections, and concatenating feature maps with upsampled features, which helps it make great progress in small object detection and has strong generalization ability. Thus we choose YOLOv3 as the backbone network to train the traffic sign detection model with transfer learning [25].

We first use the Chinese traffic-sign benchmark created by Zhu et al. [21] to pre-train a traffic sign detection model. Fig. 10 shows the types of traffic signs selected from the benchmark. Then we fine-tune the model based on a small

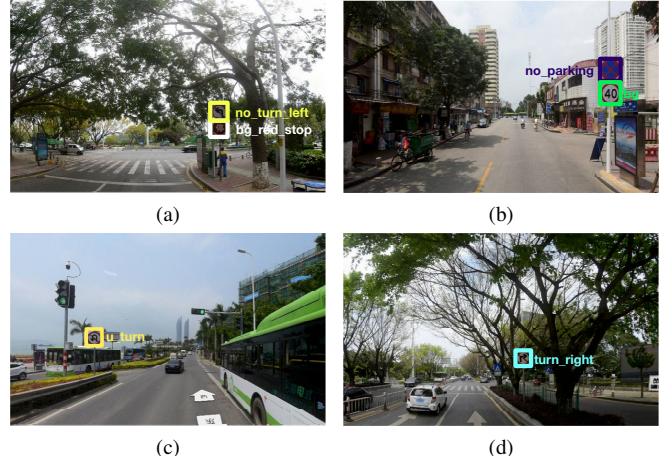


Fig. 11: Examples of training pictures. (a) A no-left-turn sign and a negative sample. (b) A no-parking sign and a negative sample. (c) A no-u-turn sign. (d) A no-right-turn sign.

dataset consisting of four categories of traffic signs, i.e., no-turn-left, no-turn-right, no-u-turn, and no-parking, collected from the targeted city. The small dataset is consistent with the street view picture dataset, as shown in Fig. 11. Therefore, if we plan to transfer the model to another city, we only need to collect a small dataset consisting of the interested categories of traffic signs from the new city and fine-tune the pre-trained model rather than collect a large dataset and re-train a new detection model.

B. Traffic Violation Identification

1) Illegal Turning: If there exists the corresponding traffic sign in the street view picture sequence before the turning behavior, such as a no-left-turn sign for a left-turn behavior, this behavior can be identified as an illegal turning behavior. We denote the no-turning road intersections in the city as $L = \{l_1, l_2, \dots, l_m\}$, where $l_i(i = 1, \dots, m)$ is the distribution of illegal turning behaviors on each unique intersection.

2) Illegal Parking: From the traffic sign database constructed above, we can get the locations of the no-parking traffic signs. After mapping these locations to the nearest road segments, we get the list of no-parking road segments. If the corresponding road segment of a parking behavior is in the no-parking road segment list and the direct distance between the no-parking sign and the parking behavior is less than a threshold ζ , the parking behavior can be identified as an illegal parking behavior. We denote the no-parking segments in the city as a set $B = \{b_1, b_2, \dots, b_n\}$, where $b_i(i = 1, \dots, n)$ is the distribution of illegal parking behaviors on each unique road segment.

3) Speeding: As for speeding, we first collect road speed restrictions from Gaode Map Open Platform¹ and get the list of speed restrictions on the roads. Then we compare the average speed with the speed restrictions of the corresponding roads. If the average speed exceeds the speed restriction on the road, it can be identified as a speeding behavior. We denote the roads with speed restrictions in the city as $S = \{s_1, s_2, \dots, s_l\}$, where l is the number of roads, $s_i(i = 1, \dots, l)$ is the distribution of speeding behaviors on each unique road.

¹<https://lbs.amap.com>

C. Traffic Violation-Prone Location Inference

After identifying traffic violations, we extract the spatiotemporal patterns of traffic violations to infer the traffic violation-prone locations.

For each traffic violation-prone location candidate, we aggregate the traffic violations related to this location. More specifically, We aggregate different illegal turning behaviors on the same intersection, parking behaviors on the same road segment and speeding behaviors on the same road.

For each traffic violation-prone location candidate r_n ($r_n \in L \cup B \cup S$), we aggregate its hourly traffic violations to build the temporal profile, i.e.,

$$\Phi(r_n) = [tv_1, tv_2, \dots, tv_H] \quad (8)$$

where tv_i ($i = 1, 2, \dots, H$) is the number of traffic violations of the i^{th} hour and H is the number of hours. Then for each traffic violation-prone location candidate, we aggregate and average hourly traffic violations in the dataset over a typical day to determine the threshold of each hour in a day to infer traffic violation-prone locations, i.e.,

$$\overline{\Phi(r_n)} = [\overline{tv_1}, \overline{tv_2}, \dots, \overline{tv_{24}}] \quad (9)$$

where \overline{tv}_i ($i = 1, 2, \dots, 24$) is the average number of aggregated traffic violations of the i^{th} hour in a day.

For each hour j ($j = 1, 2, \dots, 24$), we set the threshold of this hour as the average number of the traffic violations plus the double standard deviation of the traffic violations in the city [42], [43], [44], i.e.,

$$\mu_j = \frac{\sum_{i=1}^N \overline{tv}_j^i}{N}, \sigma_j = \sqrt{\frac{\sum_{i=1}^N (\overline{tv}_j^i - \mu_j)^2}{N}}, th_j = \mu_j + 2\sigma_j \quad (10)$$

where N is the number of traffic violation-prone location candidates, \overline{tv}_j^i is the typical traffic violation of r_i at hour j . Therefore, we can get the traffic violation-prone locations for each hour.

VII. EVALUATION

In this section, we evaluate the performance of the proposed method based on two crowd-sensed, large-scale, and real-word trajectory datasets from two cities in China, Xiamen and Chengdu, respectively². We first introduce the datasets and experiment settings. Then we present the evaluation results and a traffic violation-prone location inference system. Finally, we conduct several case studies.

A. Dataset Description

1) Taxi Trajectory Data in Xiamen and Car-Hailing Trajectory Data in Chengdu: The taxi trajectory data in Xiamen are provided by Xiamen Traffic Management Department. After a data cleansing process that removes invalid records. We obtain the taxi trajectories of 5,486 vehicles in Xiamen, Fujian Province, in China. The car-hailing trajectory data are provided

²Sample datasets and codes can be found at: <https://github.com/zhihanjiang/iTV>

TABLE I: Vehicle Trajectory Dataset Description

City	Xiamen
Duration	09/01/2016 00:00:00 - 09/30/2016 23:59:59
Record	352,300,768 (5378 vehicles)
Latitude (WGS84 ^a)	24.369406°N - 24.619351°N
Longitude (WGS84)	117.990364°E - 118.265022°E
City	Chengdu
Duration	11/01/2016 00:00:00 - 11/30/2016 23:59:59
Record	1,096,608,443 (6,096,022 orders)
Latitude (WGS84 ^a)	30.655297°N - 30.730203°N
Longitude (WGS84)	104.039652°E - 104.127062°E

^a World Geodetic System 1984, the reference system for the Global Positioning System (GPS)

TABLE II: The number of different speed restrictions on roads in Xiamen and Chengdu.

Speed Limit (km/h)	5	30	40	50	60	70	80
Xiamen	5	2	89	25	25	3	1
Chengdu	6	17	220	8	28	1	1

by GAIA Open Dataset³ from DiDiChuxing, the largest online car-hailing service provider in China, which handles around 11 million orders per day all over China⁴. After a data cleansing process that removes invalid records. We obtain the vehicle trajectories of 6,096,022 orders. These two datasets are both crowd-sensed and large-scale. The detailed summary of these two datasets are shown in TABLE. I.

2) Speed Restriction Information: We collect the speed restrictions on 150 roads in Xiamen and 281 roads in Chengdu from Gaode Map Open Platform, as shown in TALE. II.

3) Street View Pictures for Traffic Sign Detection Model Training: We drop the traffic signs the numbers of which are less than 100 and select 23923 traffic signs (45 categories) of 10267 pictures from the Chinese traffic-sign benchmark [21]. We denote this dataset as DS_1 . Moreover, we collect the small dataset of street view pictures from target cities. The small dataset consists of 162 no-left-turn signs, 127 no-right-turn signs, 108 no-u-turn signs, 106 no-parking signs and 149 negative samples, denoted as DS_2 . DS_1 and DS_2 are used to train the traffic sign detection model.

4) Street View Pictures for Driving Behavior Context Augmentation: After driver perspective modeling, 79,855 and 47,088 street view pictures are collected in Xiamen and Chengdu, respectively from Baidu Map⁵, and these street view pictures are input to the traffic sign detection model to get the traffic rule information. Specifically, the field of view is a significant parameter while collecting street view pictures. According to the basic design criteria for state highway from the New Zealand Transport Agency, when the driving speed is 0 km/h, the driver's horizontal angle of field is 180°. When the speed increases to 60 km/h, the angle decreases to 74°, and when the speed increases to 80 km/h and 100 km/h, the angle decreases to 60° and 40°, respectively. Generally, the

³<https://outreach.didichuxing.com/research/opendata/>

⁴<https://www.didiglobal.com>

⁵<https://map.baidu.com>

TABLE III: The Evaluation Results on Traffic Sign Detection

Signs	AB	FC	CNNs	Proposed
No-left-turn	0.375	0.563	0.563	0.688
No-right-turn	0.462	0.693	0.538	0.846
No-u-turn	0.727	0.455	0.545	0.818
No-parking	0.182	0.636	0.818	0.818
Total (mAP)	0.437	0.587	0.616	0.793

speed of the vehicle on the urban road is lower than 60 km/h. Thus, the field of view used in street view collection is 74°.

B. Experiment Settings

1) *Evaluation Plan*: In traffic sign detection, we first separate DS_1 into training set and validation set, 80% and 20% respectively, and train a traffic sign detection model based on the YOLOv3 [24] model. Then from DS_2 , we select 10% of each type of traffic signs as the test set and we fine turn the model with the rest of the traffic signs, which is separated into training set and validation set, 80% and 20% respectively. The traffic sign detection model is used to detect traffic signs in 79,855 and 47,088 street view pictures collected from Xiamen and Chengdu, respectively. After that, we get the traffic violation distribution so as to infer the traffic violation-prone locations. In traffic violation-prone location inference, we show the thresholds of traffic violation-prone locations and the number of traffic violation-prone locations in Chengdu and Xiamen during a month, respectively. Finally we evaluate the runtime performance, present a traffic violation-prone location inference system, and give several case studies.

2) *Evaluation Metric*: We evaluate the performance of the traffic sign detection model on the test set using the average precision of interested categories of traffic signs, i.e., $AP = \sum_{k=1}^N p(k)(r(k-1) - r(k))$, where $p(k)$ and $r(k)$ are the precision and recall at the k th threshold, and mAP is the mean of AP for all categories.

3) *Baseline Methods*: We compare our method with following traffic sign detection models.

- **AB** [17]: This method uses a set of Haar wavelet features obtained from AdaBoost training to detect traffic signs, and Bayesian generative modeling is used to classify traffic signs.
- **FC** [20]: This method consists of two deep learning components including fully convolutional network (FCN) guided traffic sign proposals and deep convolutional neural network (CNN) for object classification.
- **CNNs** [21]: This method uses a end-to-end convolutional neural network (CNN) to detect and classify traffic signs.

C. Traffic Sign Detection Results

The evaluation results on traffic sign detection are shown in TABLE. III. The street view pictures are taken in complicated urban environment with interferences and have low resolution. Besides, the size of traffic signs in the pictures are varied, and since the street view pictures are panoramic and the shooting angles are different, the edges of pictures are usually kind of distorted, which increases the difficulties of traffic sign detection. According to the evaluation results, the AB, FC,

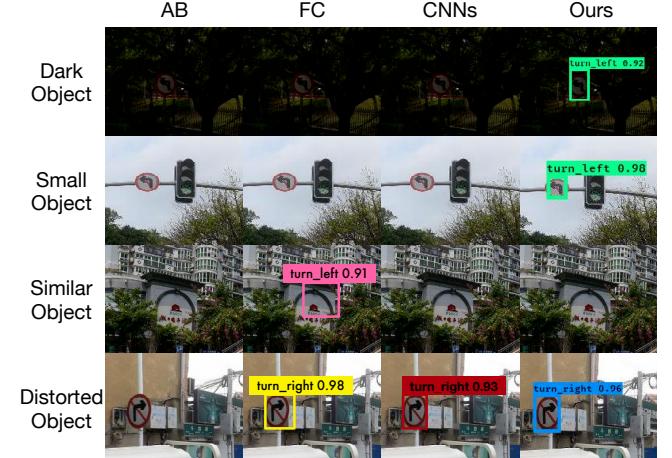


Fig. 12: The performance of traffic sign detection models under severe conditions.

TABLE IV: Summary of Traffic signs Detected in Xiamen and Chengdu

City	No-left-turn	No-right-turn	No-u-turn	No-parking	Total
Xiamen	282	110	29	848	1269
Chengdu	198	71	19	344	632



Fig. 13: Some examples of traffic sign detected in street view pictures.

and CNNs methods do not perform well on our test set, and the fine-tuned YOLOv3 model in our proposed framework achieves the best performance. Furthermore, Fig. 12 shows examples of traffic sign detection under severe conditions. The AB method does not perform well when the traffic sign is in dark environment, very small, or distorted, but it rarely misidentify similar objects as traffic signs, while the FC method recognizes similar objects in pictures as traffic signs more frequently. The CNNs method works well when traffic signs are clear and it rarely misidentify similar object, but it fails when traffic signs are blurred. Different from the baseline methods, the proposed fine-tuned YOLOv3 model works well under these severe conditions.

Therefore, we use the fine-tuned YOLOv3-based traffic sign detection model to detect the traffic signs in 79,855 and 47,088 street view pictures collected from Xiamen and Chengdu, respectively. The summary of the traffic signs detected from the street view pictures in Xiamen and Chengdu are shown in TABLE. IV and Fig. 13 shows some examples of traffic signs detected in street view pictures.

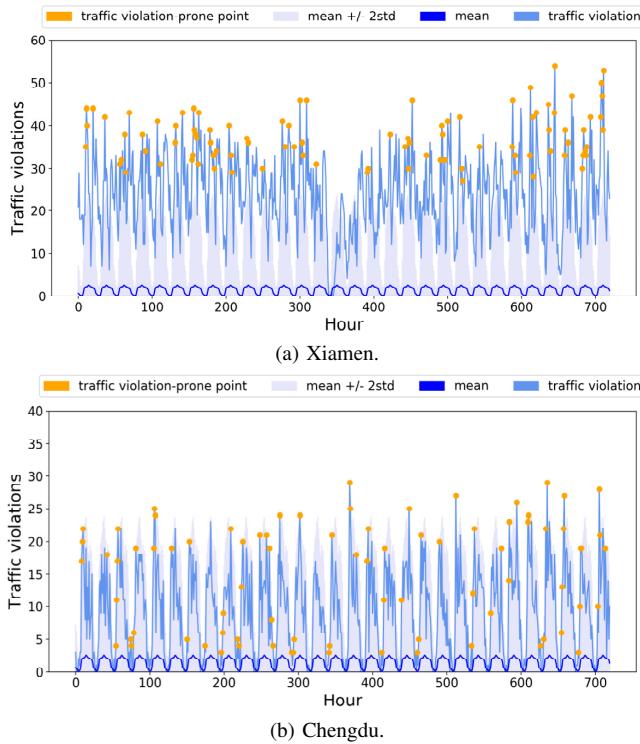


Fig. 14: The thresholds of traffic violation-prone locations in Xiamen or Chengdu, and the points in each city represent the inferred traffic violation-prone points of a location in the city in a month. 'std' means standard deviation.

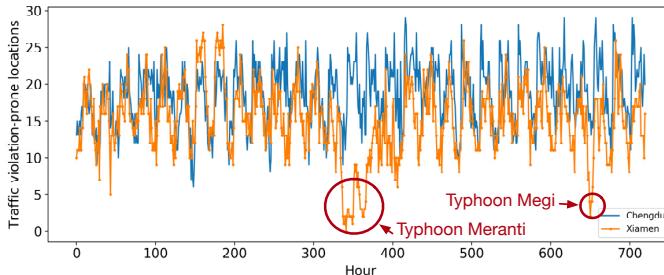


Fig. 15: The number of traffic violation-prone locations in Chengdu (November) and Xiamen (September) in 2016.

D. Traffic Violation-Prone Location Inference Results

Fig. 14 shows the thresholds of traffic violation-prone locations in Xiamen and Chengdu, respectively, and the points marked in the figure represent the inferred traffic violation-prone points of a location in the corresponding city in a month. Fig. 15 shows the number of traffic violation-prone locations in Chengdu, November, 2016, and Xiamen, September, 2016, respectively. Specifically, there were two typhoons influenced Xiamen in this month, Meranti on September 15th and Megi on September 27th, which are corresponding to the lowest two segments on the curve of Xiamen.

E. Runtime Performance

The typical temporal profiles of traffic violation-prone location candidates are updated every day when data from a new day have been collected. We deploy our system on a server with NVIDIA GeForce GTX 1080 Ti 11GB and 32GB RAM, and it takes an average of 55.937 ms and 0.209 ms to do traffic sign detection for one picture and map matching for

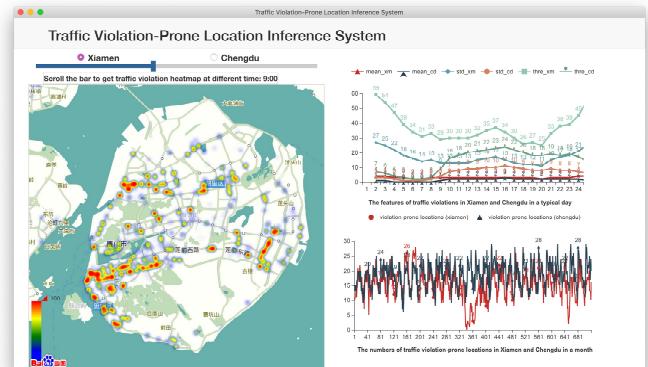


Fig. 16: Traffic Violation-Prone Location Inference System.



Fig. 17: An illegal turning-prone location in Chengdu. (a) The street view picture of an illegal turning-prone location. (b) An illegal turning-prone location

one GPS point, respectively. In Xiamen, the average number of GPS points and street view pictures that need to be processed every day were about 11,743,359 and 2,662. In Chengdu, these numbers are 36,553,615 and 1,570, respectively. Therefore, the average time of processing new data collected every day is about 43.38 minutes in Xiamen and 128.79 minutes in Chengdu.

F. Traffic Violation-Prone Location Inference System

We build a traffic violation-prone location inference system⁶. From this system, we can easily find the traffic violation-prone locations in a city at different time, as shown in Fig. 16.

G. Case Studies

1) *Illegal Turning*: Illegal turning is the second most frequent traffic violation in Chengdu, and we found a lot of illegal turning behaviors happen on the intersection of Wucheng Street and Dongan South Road, as shown in Fig. 17. In this intersection, drivers are forbidden to turn left (the red arrow in the figure) after crossing a bridge, but there are still many people violating the traffic regulation.

2) *Illegal Parking*: According to the Xiamen Traffic Police, illegal parking is the most frequent traffic violation from 2015 to 2018. Through observing heat maps, we can find that the intersection of Chenggong Avenue and Nanshan Road had a large number of illegal parking behaviors. However, after an in-depth investigation, we found that a new subway station was being built during that time, which led to terrible traffic conditions there. Thus there were many illegal parking behaviors. Besides, Xiahe Road, Siming South Road, and many

⁶The system can be visited at: <http://zhihanjiang.com/tv-infer-web/>



Fig. 18: An illegal parking-prone location in Xiamen. (a) The street view picture of an illegal parking-prone location in Xiamen. (b) An illegal parking-prone location.



Fig. 19: A speeding-prone location in Xiamen. (a) The street view picture of a speeding-prone location. (b) A speeding-prone location.

other roads near the tourist hotspots are also traffic violation-prone locations. Siming South Road is a very busy road in Xiamen, along which there are a lot of tourist hotspots, such as Xiamen University, Nanputuo Temple, Overseas Chinese Museum, and it also intersects Zhongshan Road, which is a famous tourist road. The complicated environment and the large traffic volume make it become the road with maximum illegal parking behaviors in Xiamen. As shown in Fig. 18, some vehicles are parked on the road, although there is a no-parking sign erected there.

3) *Speeding*: Speeding is the sixth most frequent traffic violation in Xiamen. Fig. 19 shows a speeding behavior on Bailuzhou Road in Xiamen. Bailuzhou Road is a speeding-prone location in Xiamen, which is a 4-lane dual carriageway. The speed restriction of Bailuzhou Road is 60 km/h. Besides we found a lot of speeding behaviors on Jiahe Road and Chenggong Avenue. They are both important roads crossing over Xiamen Island.

VIII. CONCLUSION

In this work, we propose a low-cost, comprehensive and dynamic method for inferring the traffic violation-prone locations in cities based on the large-scale vehicle trajectory data and road environment data to provide some insights for the traffic management department about traffic dynamics in cities to help optimize the utility and effectiveness of the traffic enforcement strategies. Firstly, we normalize the trajectory data by map matching algorithms and get the driving behavior distribution. Secondly, we model match driving behaviors to corresponding road segments and restore the spatiotemporal contexts of driving behaviors to get the traffic rule information so that we can get three types of traffic violation distributions, i.e., illegal turning, illegal parking and speeding. Then we extract the spatiotemporal patterns of traffic violations to infer the traffic violation-prone locations in cities. Finally, we build

a traffic violation-prone location inference system and give some case studies. The proposed method is evaluated using crowd-sensed, large-scale, and real-world datasets.

In the future, we plan to broaden and deepen this work in two directions. Firstly, we plan to incorporate more trajectory open data sources from other cities. Secondly, we plan to incorporate urban environment data, such as Points of Interests and traffic volumes, to explore more in-depth relationships between traffic violation-prone locations and the urban environment.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their constructive suggestions. This research is supported by NSF of China (61802325, 61872306 and U1605254), NSF of Fujian Province (2018J01105), the China Fundamental Research Funds for the Central Universities (20720170040), and Xiamen Science and Technology Bureau (3502Z20193017).

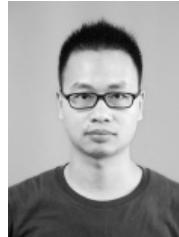
REFERENCES

- [1] G. Zhang, K. K. Yau, and G. Chen, "Risk factors associated with traffic violations and accident severity in China," *Accident Analysis & Prevention*, vol. 59, pp. 18–25, 2013.
- [2] V. Coric and M. Gruteser, "Crowdsensing maps of on-street parking spaces," in *2013 IEEE International Conference on Distributed Computing in Sensor Systems*. IEEE, 2013, pp. 115–122.
- [3] E. Moyano-Díaz, "Evaluation of traffic violation behaviors and the causal attribution of accidents in Chile," *Environment and Behavior*, vol. 29, no. 2, pp. 264–282, 1997.
- [4] M. O. o. O. New York City, "Vision Zero Year Five Report," New York, Tech. Rep., 2019.
- [5] T. Alternatives, "From Chaos to Compliance: How the NYPD Can Grasp New York City's Traffic Safety Problem," New York, Tech. Rep., Aug. 2009.
- [6] C. Chen, D. Zhang, X. Ma, B. Guo, L. Wang, Y. Wang, and E. Sha, "Crowddeliver: Planning city-wide package delivery paths leveraging the crowd of taxis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1478–1496, 2016.
- [7] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining Interesting Locations and Travel Sequences from GPS Trajectories," in *Proc. WWW 2009*. ACM Press, 2009, pp. 791–800.
- [8] L. O. Alvares, V. Bogorni, B. Kuijpers, J. A. F. de Macedo, B. Moelans, and A. Vaisman, "A model for enriching trajectories with semantic geographical information," in *Proc. GIS 2007*. ACM Press, 2007, pp. 22–30.
- [9] D. Zhang, N. Li, Z.-H. Zhou, C. Chen, L. Sun, and S. Li, "iBAT: Detecting anomalous taxi trajectories from GPS traces," in *Proc. UbiComp 2011*. ACM Press, 2011, pp. 99–108.
- [10] L. Chen, D. Zhang, G. Pan, X. Ma, D. Yang, K. Kushlev, W. Zhang, and S. Li, "Bike Sharing Station Placement Leveraging Heterogeneous Urban Open Data," in *Proc. UbiComp 2015*. ACM, 2015, pp. 571–575.
- [11] F. Zhang, D. Wilkie, Y. Zheng, and X. Xie, "Sensing the Pulse of Urban Refueling Behavior," in *Proc. UbiComp 2013*. ACM Press, 2013, pp. 13–22.
- [12] L. Chen, D. Zhang, X. Ma, L. Wang, S. Li, Z. Wu, and G. Pan, "Container port performance measurement and comparison leveraging ship GPS traces and maritime open data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1227–1242, 2016.
- [13] T. He, J. Bao, R. Li, S. Ruan, Y. Li, C. Tian, and Y. Zheng, "Detecting Vehicle Illegal Parking Events using Sharing Bikes' Trajectories," *SIGKDD*. ACM, 2018.
- [14] J. T. Lee, M. S. Ryoo, M. Riley, and J. Aggarwal, "Real-time illegal parking detection in outdoor environments using 1-D transformation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 1014–1024, 2009.
- [15] E. S. Gerber, "Vehicle speeding detection and identification," Jan. 1995, uS Patent 5,381,155.
- [16] A. De La Escalera, L. E. Moreno, M. A. Salichs, and J. M. Armingol, "Road traffic sign detection and classification," *IEEE transactions on industrial electronics*, vol. 44, no. 6, pp. 848–859, 1997.

- [17] C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler, "A system for traffic sign detection, tracking, and recognition using color, shape, and motion information," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005*. IEEE, 2005, pp. 255–260.
- [18] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1484–1497, 2012.
- [19] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The German Traffic Sign Detection Benchmark," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2013, pp. 1–8.
- [20] Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, and W. Liu, "Traffic sign detection and recognition using fully convolutional network guided proposals," *Neurocomputing*, vol. 214, pp. 758–766, 2016.
- [21] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-Sign Detection and Classification in the Wild," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [22] H. Luo, Y. Yang, B. Tong, F. Wu, and B. Fan, "Traffic sign recognition using a multi-task convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1100–1111, 2018.
- [23] W. Huang, M. Huang, and Y. Zhang, "Detection of Traffic Signs Based on Combination of GAN and Faster-RCNN," in *Journal of Physics: Conference Series*, vol. 1069, no. 1. IOP Publishing, 2018, p. 012159.
- [24] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [25] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
- [26] P. Newson and J. Krumm, "Hidden Markov map matching through noise and sparseness," in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2009, pp. 336–343.
- [27] M. Ren and H. A. Karimi, "A hidden Markov model-based map-matching algorithm for wheelchair navigation," *The Journal of Navigation*, vol. 62, no. 3, pp. 383–395, 2009.
- [28] D. Bernstein, A. Kornhauser *et al.*, "An introduction to map matching for personal navigation assistants," 1996.
- [29] C. E. White, D. Bernstein, and A. L. Kornhauser, "Some map matching algorithms for personal navigation assistants," *Transportation research part c: emerging technologies*, vol. 8, no. 1–6, pp. 91–108, 2000.
- [30] C. Chen, Y. Ding, X. Xie, S. Zhang, Z. Wang, and L. Feng, "Trajcompressor: An online map-matching-based trajectory compression framework leveraging vehicle heading direction and change," *IEEE Transactions on Intelligent Transportation Systems*, 2019.
- [31] D. Vaughn, "Vehicle speed control based on GPS/MAP matching of posted speeds," Jan. 1996, uS Patent 5,485,161.
- [32] J. S. Greenfeld, "Matching gps observations to locations on a digital map," in *81th annual meeting of the transportation research board*, vol. 1, no. 3. Washington, DC, 2002, pp. 164–173.
- [33] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk, "On map-matching vehicle tracking data," in *Proceedings of the 31st international conference on Very large data bases*, 2005, pp. 853–864.
- [34] A. Dewandaru, A. M. Said, and A. N. Matori, "A novel map-matching algorithm to improve vehicle tracking system accuracy," in *2007 International Conference on Intelligent and Advanced Systems*. IEEE, 2007, pp. 177–181.
- [35] A. Harati-Mokhtari, A. Wall, P. Brooks, and J. Wang, "Automatic Identification System (AIS): Data reliability and human error implications," *The Journal of Navigation*, vol. 60, no. 3, pp. 373–389, 2007.
- [36] F. van Diggelen, *GNSS Accuracy, Lies, Damn Lies, and Statistics [on-Line]*, 2007.
- [37] U. Gather and V. Schultze, "Robust estimation of scale of an exponential distribution," *Statistica Neerlandica*, vol. 53, no. 3, pp. 327–341, 1999.
- [38] G. D. Forney, "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.
- [39] J. H. Williams, *Quantifying Measurement*. Morgan & Claypool Publishers, 2016.
- [40] C. L. Lawson and R. J. Hanson, *Solving least squares problems*. Siam, 1995, vol. 15.
- [41] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7263–7271.
- [42] F. Pukelsheim, "The three sigma rule," *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994.
- [43] D. J. Wheeler, D. S. Chambers *et al.*, *Understanding statistical process control*. SPC press, 1992.
- [44] V. Czitrom and P. D. Spagon, *Statistical case studies for industrial process improvement*. Siam, 1997, vol. 1.



Zhihan Jiang received the B.Sc. degree in computer science and technology from Xiamen University, China in 2018. She is currently pursuing the master's degree in Department of Computer Science, Xiamen University.



Longbiao Chen is an assistant professor with Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University. He obtained his Ph.D. degree in computer science from Sorbonne University, France in 2018 and Zhejiang University, China in 2016. His research interests include Ubiquitous Computing, Crowd Computing, and Urban Computing.



Binbin Zhou received her M.Phil degree in computer science from Hongkong Polytechnic University, in 2011. She is currently pursuing a Ph.D. degree in the Department of Computer Science, Zhejiang University. Her research interests are mainly in urban computing, big data applications, and ubiquitous computing.



Jinchun Huang received the B.Sc. degree in computer science and technology from Xiamen University, China in 2020.



Tianqi Xie received the B.Sc. degree in computer science and technology from Xiamen University, China in 2020.



Xiaoliang Fan is a senior research engineer with Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University. He received his Ph.D. degree at University Pierre and Marie CURIE, France in 2012. His research interests include services computing, and big data analytics.



Cheng Wang received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, China in 2002. He is a Professor with Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, China. His research interests include remote sensing image processing, mobile LiDAR data analysis, and multisensor fusion.