

Inferring and Analysis Drivers Violation Behavior Through Trajectory

Zouqing Cai, Wei Pei, Yongying Zhu, Mingyu Lu, Lei Wu

▶ **To cite this version:**

Zouqing Cai, Wei Pei, Yongying Zhu, Mingyu Lu, Lei Wu. Inferring and Analysis Drivers Violation Behavior Through Trajectory. 2nd International Conference on Intelligence Science (ICIS), Oct 2017, Shanghai, China. pp.295-304, 10.1007/978-3-319-68121-4_32 . hal-01820920

HAL Id: hal-01820920

https://inria.hal.science/hal-01820920

Submitted on 22 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

Inferring and Analysis Drivers Violation Behavior through Trajectory

Zouqing Cai, Wei Pei, Yongying Zhu,
Mingyu Lu, Lei Wu,

Information Science and Technology College,
Dalian Maritime University,
Dalian, Liaoning Province, China,
{Zouqing.Cai, Wei.Pei, Yongying.Zhu, Mingyu.Lu,
Lei.Wu,peiweidl}@gmail.com

Abstract. In this paper, we present an algorithm for inferring violation movements and categorizing levels of driving behavior. With this algorithm we extract the speeding and retrograde behavior from the real trajectories datasets of Xinjiang, analyze the changing regulation of six streets in the working day and day off on the overall and explore the driving characteristics which is very dangerous（level 4）. The results of this study can not only be used for early warning of drivers violations, but also provide the data support and decision basis for the traffic management department to master the situation of traffic violations and formulate the traffic management counter measures.

Keywords: trajectory data mining; violation driving behavior; violation movements; driver classification

1 Introduction

The global status report on road safety 2015 indicates that worldwide the total number of road traffic deaths has plateaued at 1.25 million per year [1]. In 2015, the number of all kinds of production safety accident deaths has plateaued at 66182 in our country and 54% of the deaths are caused by traffic accidents. Drunk driving, speeding, retrograde, running the red light and other illegal operations account for more than 60% of the proportion of traffic accidents [2]. Due to a rising trend in vicious traffic accidents, the study of violation driving behavior has been a hot spot in recent years.

Several works have been proposed for driving behavior analysis in simulation systems [4-8]. For instance, sensors and simulators are used to recognize the fatigue driving state [6].

With the popularity of mobile devices such as GPS and cell phones, large amounts of real driving traces are available to analyze the behavior of drivers [9-14].

By inferring violation behavior and sum up the law of the illegal driving behavior, the traffic management departments can grasp the situation of traffic violations and

take effective measures to reduce or even avoid major traffic accidents. At the same time, it has a warning effect on the drivers who often break the rules. In this paper, we focus on real trajectories of drivers, propose an efficient algorithm to identify the violation behaviors of drivers based on individual trajectories and classify the drivers according to the risk weights. According to the different function area, six roads including four one-way streets are picked up to demonstrate the algorithm. The first two steps of the algorithm extract the speeding and retrograde behavior, analyze the change regulation of illegal driving behaviors either workday or rest day and discuss the place where the driver often has retrograde behaviors. Furthermore the violation behavior of drivers is evaluated with the last two steps of the algorithm and the very dangerous drivers'（level 4） personality characteristics is analyzed.

The rest of the paper is organized as follows. Section2 introduces the related work in illegal driving behavior analysis. Section3 presents the main definitions used in this paper, experiments with real data and proposes a method to find speeding and retrograde behavior. Section4 shows and discusses the experiments results and Section5 concludes the paper.

2 RELATED WORK

Several works have been done to study the driver behaviors in simulation systems. Kedar-Dongarkar and Das presented a new method of driver classification for optimizing energy usage [5]. Vehicle acceleration, braking, speed, and throttle pedal were used to classify the drivers into three categories: aggressive, moderate and conservative. Hu presented a new identification method of fatigue driving state, which is obtained from driving behavior data analysis both about normal driving state and the fatigues one [6]. Quantitative evaluation of driving styles by normalizing driving behavior was studied in [7].

Some works which analyze driving behavior using real GPS trajectories will be introduced as follows. For example, Carboni was interested in abnormal behaviors of individual trajectories of drivers, and presented an algorithm for finding anomalous movements and categorizing levels of driving behavior [10]. Li proposed a method to realize the safety analysis of vehicle driving behavior by using GPS vehicle trajectory data [11]. They studied the proportion of the bus over-speed time in entire driving time. Dueholm J V presented a method of latent semantic information mining for trajectory data [12]. He M proposed WhozDriving to solve the issues of detecting abnormal driving trajectory in driver's history trajectories[13]. Chen and Zhang focused on abnormal trajectories of taxis that deviated the standard route from origin and destination, where the standard route represented the path followed by the majority of taxis [14].

Although the previous detailed works analyze several illegal characteristics of driving, most of them have not been developed with real trajectories. Some works used the real trajectory analysis of dangerous driving behavior, but they concentrated on how to judge the abnormal behavior such as rapid change, sharp turn speeding etc. They rarely conduct a comprehensive analysis of illegal behavior and summarize the rules of violation driving. In this paper we focus on real trajectories of drivers,

propose an algorithm to identify violation behavior and classify drivers into different levels of danger. The effectiveness of the algorithm is greatly proved by experiments.

3 Discover Traffic Violation

In this section we firstly present some basic definitions and formulas (Section3.1). Secondly, filter trajectories will be introduced in Section3.2. Finally, an algorithm will be proposed to find and evaluate violation driving behavior in Section3.3.

3.1 Main Definitions

The basic definitions for trajectories are described as follows.

Definition 1. Point. A point is defined as p =(id, x ,y ,t ,s ,h), where id is a vehicle ID represents a driver, x and y are the latitude and longitude that represent space ,t is the *timestamp* in which the point has been collected and h is a direction angle that represents the driving direction.

Definition 2. Trajectory. A trajectory is expressed as T={r_1, $r_2 \ldots r_n$}, where r_i={$p_1, p_2 \ldots p_n$}, p_i=(id, x_i, y_i, t_i, s_i, d_i), $t_1<t_2<\ldots<t_n$.A trajectory includes one or several roads, and r represents a road.

Definition 3. Subtrajectory. A subjectory s of T is a series of points $<p_k, p_{k+1}$,…,$<p_j$, where $p_k \subset T$ and 1≤k≤j and j≤n. (In this paper, subtrajectory division is based on road. A road of trajectory is corresponds to a subtrajectory.)

Definition4. Trajectories. A trajectories is T_s ={ T_1 , T_2 … T_n }, where T_i={ $p_1, p_2 \ldots p_n$}, pi=(id,x_i,y_i,t_i,s_i,d_i) , $t_1<t_2<\ldots<t_n$.

Definition 5. Violation movement. A trajectory has violation movements when it has at least one subtrajectory with violation s such as the driving speed exceeds the road speed limit, the driving direction is contrary to the normal driving direction.

The case which meets the equation (1) is judged as the over speed and the one meeting the equation (2) is retrograde. v is the speed of the point, and h stands for the direction of the point. r represents a road ,and r=1 presents a one-way street. v' represents the maximum value of road speed limit, and h' is the direction of road restrictions.

$$c = \begin{cases} 1, & v > v' \\ 0, & v \le v' \end{cases} \quad (1) \qquad d = \begin{cases} 1, r = 1 \cap |h - h'| \le 45 \\ 0, \ other \end{cases} \quad (2)$$

After defining the violation movement, the movements are analyzed deeply. Some characteristics are found to evaluate driving behavior. In this analysis we mainly consider three features:

F1: The driver only has a violation movement in a road.

F2: Continuous illegal behaviors in a road.

F3: Continuous illegal behaviors in different roads.

Based on these three features, we defined four levels of drivers:

Level 1(Careful driver): A careful driver is without any violation behavior. Although someone may complain that it makes no sense to discover careful drivers, it is very useful for the company to give a reward or compliment to the good drivers.

Level 2(Distract driver): A distract driver is with feature F1. The driver is distracted while driving with the performance of many secondary tasks, including texting and dialing cell phones and so on.

Level 3(Dangerous driver): A dangerous driver is someone with feature F2. Such as a driver continues to over-speed on the road.

Level 4(Very dangerous driver): A very dangerous driver is someone with features F3. For example, the driver has a continuous speeding or retrograde driving behavior in different roads.

3.2 Filter Trajectories

We get real trajectories of taxis collected at February 2016 in the city of Xinjiang, China. The dataset has about 300 million data with points collected at intervals of 30 second, including vehicleid, gettime, storetime, speed, direction, latitude and longitude. The taxi dataset is for a city scale but we just focus on several specific routes, so we implement filter trajectories which has the same origin and the same destination on a road [13]. Using this method, trajectories are filtered in six routes where close to different function area.

3.3 Discover Speeding Or Retrograde

In this paper an algorithm is proposed to discover violation driving behaviors. Firstly it identifies violation movements based on over-speed and retrograde driving behavior. Secondly it evaluates the driving behavior according to the formulas which we define.

Because the violations driving behavior is very transitory, the subtrajectories with violations behavior are normally only a few points. If considering the violation movements of each point, noise can be introduced. If considering too many points (three or more) the violation movement may not be captured. So after some analysis and experiments on real trajectory data, it consider that at least two consecutive points should have violation movements for a subtrajectory to characterise the violation behavior. Violation behavior can be captured well for trajectories with frequently sampled points, like 1 or 2 seconds. A dataset with sampling rate as 30 seconds, for instance, can also reveal violation behavior.mm.

The algorithm of findViolation is as follows.

findViolation:

Ts=$\{t_1, t_2 \ldots t_n\}$*.For each trajectory* t_i*, using the length of the road to overlap violations,* $t_i=\{r_1, r_2 \ldots r_n\}$*, vio(*$t_i$*)=vio(*$r_1$*)+vio(*$r_2$*)+…+vio(*$r_n$*).*

$r_i=\{p_1, p_2 \ldots p_n\}$*. For each point ,* c_i *and* d_i *can be computed according to equation (1)and(2).*

The count of subtrajectories violations movments can be computed according to the road ..

$$\text{violation}(r_i)= \sum(c_i * c_{i+1} + d_i * d_{i+1}) \quad (3)$$

The violation behavior of drivers is evaluated using formulal(4)

$$level = \begin{cases} level1, vio(t_i) = 0 \\ level2, vio(r_i) = 1 \ and \ vio(r_i) + \cdots + vio(r_{i-1}) + vio(r_{i+1}) + \cdots + vio(r_n) = \\ level3, 4 > vio(r_i) > 1 \ and \ vio(r_1) + \cdots + vio(r_{i-1}) + vio(r_{i+1}) + \cdots + vio(\\ level4, vio(r_i) > 1 \ and \ vio(r_1) + \cdots + vio(r_{i-1}) + vio(r_{i+1}) + \cdots + vio(r_n) > \\ \quad or \ vio(r_i) > 4 \ and \ vio(r_1) + \cdots + vio(r_{i-1}) + vio(r_{i+1}) + \cdots + vio(r_n) = \end{cases} \quad (4)$$

In findViolation, the first two steps are used to discover the violation movements, which named fv-1. The last two steps of the algorithm are used to divide the drivers into different levels of violation, which named fv-2.

4 Experiental Results and Dicussion

In this section our algorithm is demonstrated and the experimental results are analyzed with the real-world GPS trajectory dataset. The real dataset is processed in Section4.1. We analyzed the different change laws of different roads in working days and day off in Section4.2. The violation behavior of drivers is evaluated and the characteristics of the very dangerous driver are analyzed in Section4.3.

4.1 Data Processing

The dataset that mentioned in Section4.2 is used in the experiment. Firstly, the error points are removed. Such as the latitude or the longitude is null, or the value of speed is less than zero. Secondly, a day is divided into 24 time periods. Thirdly, all the tracks are adapted to the spatial division and established the index to speed up the search. Finally, the trajectory points are matched with the map using AMAP API.

4.2 Part one

The superscript numeral used to refer to a footnote appears in the text either directly after the word to be discussed or – in relation to a phrase or a sentence – following the punctuation mark (comma, semicolon, or period). Footnotes should appear at the bottom of the normal text area, with a line of about 5cm set immediately above them .

In this section we use February 1st, 2016 (Monday) as the example of working day and February 27th, 2016(Saturday) as the example of day off to explore the temporal and spatial evolution of speeding and retrograde on the same roads. We pick up a region which has a large number of taxis. Fig. 1 shows the heatmap of the area on Monday. Based on the color change, it is clear to show where the number of operation taxis is the largest. According to the heatmap, we choose six busy roads and use the method which is mentioned in section 3.2 to filter trajectories. The relevant information of the roads is shown in Table 1. As we all known that roads can cross over different functional areas. In order to expediently discuss the relationship between the distribution of the taxi and the different functional areas, the roads are

divided into four groups according the main utilization of land along the roads. For example, although there are two top grade residences nearby Zhongshan Rd, it is defined as a commercial road because there are several supermarkets and shopping malls on both sides of the road. Table 2 shows the results of the division.

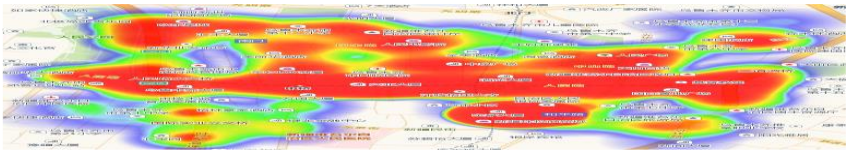


Fig. 1. The thermodynamic chart of the area.

Table 1. THE INFORMATION OF THE SIX STREETS

Detail	Renmin Rd	Democrac-y Rd	Liberati-on Rd	Zhongshan Rd	Xingfu Rd	Jianshe Rd
Rd level	main rd	main rd	main rd	main rd	main rd	secondary rd
Speed	60km/h	50km/h	50km/h	50km/h	50km/h	50km/h
Direction	null	West-east	null	East-west	Northwest-Southeast	Southeast-Northwest

TABLE 2. FONT SIZES OF HEADINGS. TABLE CAPTIONS SHOULD ALWAYS BE POSITIONED ABOVE THE TABLES.

Character	Streets
Commercial	Zhongshan Rd, Democracy Rd
Health and Education	Liberation Rd, Xingfu Rd
Administrative office	Renmin Rd
Residence Community	Jianshe Rd

Fig. 2 and Fig. 3 show the changing laws of the operation taxis in February 1st, 2016 (workday) and February 27st, 2016 (rest day). The changing rules of operating taxis during the two days have similar characteristics. For example, the number of the operation taxis gradually decreases from 00:00:00 to 07:00:00 and there are lots of taxis from 10:00:00 to 17:00:00. The distribution of operating vehicles between weekdays and rest days are different. For example, there are more operation taxis on workday than the rest day, the rest day only has a low peak and the proportion of vehicles in the night are bigger. From the micro perspective, Zhongshan Rd has a large amount of vehicles at 10:00:00-12:00:00 and 13:00:00-17:00:00 on Saturday which is related to the commercial character of this road .The character of Jianshe Rd is a residence community. The variation of the number of vehicles on Monday is quite similar to the one of Saturday, which has more vehicles on Monday at 10:00:00-12:00:00 and 16:00:00-18:00:00. This period is the peak of commuting. By

comparing the trends of the two graphs, we find a strange phenomenon that the Renmin Rd on Monday at 12: 00: 00-14:00:00 has a peak valley but on Saturday is a trough. The reason for it is that there are many office buildings near the Renmin Rd, and more staff go out to eat or go home at the lunch time of 12: 00: 00-14: 00: 00 on Monday (working day) etc.

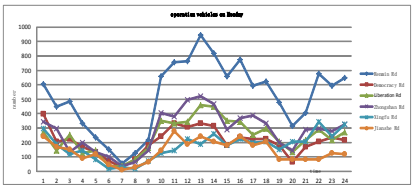


Fig. 2. The regulation of taxis on Monday.

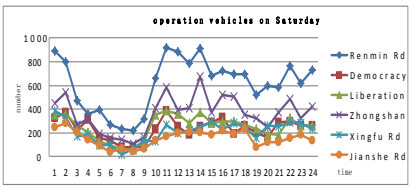


Fig. 3. The regulation of taxis on Monday.

The retrograde behavior is extracted from the real dataset by fv-1. Fig. 4 and Fig. 5 show the changing law of the retrograde vehicles. In Fig. 4 we find that there are many vehicles with retrograde behavior from 10:00:00 to 19:00:00. The number of the retrograde taxis gradually decreases from 00:00:00 to 07:00:00 because of the reducing total number of vehicles. Comparing the trend of each road in Fig. 2 and Fig. 4 we find that the general trend of illegal vehicles and operating vehicles is similar. Of course, it also has individual differences. For example, On Monday, the overall trend of operating vehicles in Zhongshan Rd from 12: 00: 00 to 14: 00: 00 is a peak valley and reaches its maximum at 13:00:00, but the trend of retrograde vehicles is a trough and drops to the lowest at 13:00:00. Owing to the trips of day off are random, the retrograde vehicles are also distributed randomly in Fig. 5.

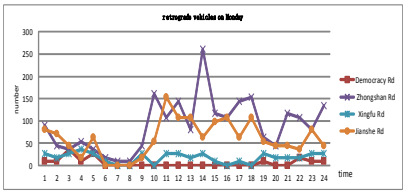


Fig. 4. The retrograde taxis on Monday.

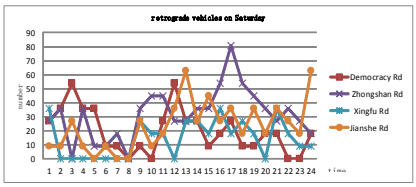


Fig. 5. The retrograde taxis on Saturday.

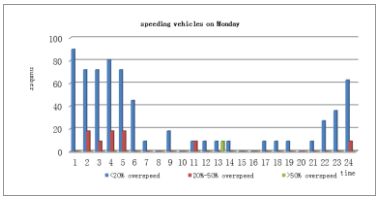
The statistics show that in addition to Xignfu Rd, the numbers of retrograde vehicles of the other three roads in the working days are more than the rest day. On Xingfu Rd there are 234 drivers with retrograde behavior on working day and there are 342 on rest day. The Xingfu Rd is used for Health and Education. There are a primary school, a secondary school and a vocational college near the road. The reason for the number of violation vehicles in the working day less than the rest day is that the students and their parents can comply with the traffic rules. Fig. 6 shows the heatmap of retrograde vehicles. We find that retrograde behavior occurs at the

crossroads with a higher frequency. So it suggest that the traffic management department should increase the supervision on the intersection.

The speeding behavior is extracted from the real dataset by the fv-1. Through analyzing Fig. 7 and Fig. 8 we find that the number of over-speed vehicles at night is higher than the one at the daytime. The number of serious over-speed vehicles in workday is more than that in the rest day. At 9: 00: 00-10: 00: 00 the number of speeding vehicles is very small. According to Fig. 2 and Fig. 3 we find that the traffic flow during this period is very large, so we speculate that in the absence of speeding vehicles during the period, the road may be in a congested state. The low-level over-speed behavior that no more than 20% is the top of the over-speed ranking .Severe speeding more than 50% is the least.



Fig. 6. The heat map of retrograde taxis on Monday.



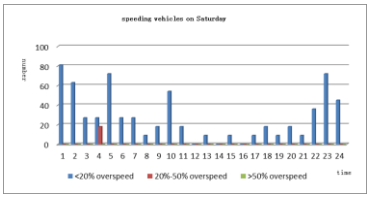


Fig. 7. The level of speeding taxis on Monday. **Fig. 8.** The level of speeding taxis on Saturday.

4.3 Part two

In this section 20 vehicles trajectories are picked up randomly from Section4.2 to find violation behaviors of individual trajectories of drivers and categorize levels of driving behavior with the algorithm. The results are shown in Table 3. In order to illustrate the violations trajectories in detail, we show a trajectory of a very dangerous driver.

Table 3. THE RESULTS OF PART TWO

Level	Results
Precise driver	15
Imprecise driver	2
Dangerous driver	0
Very dangerous driver	1

Fig. 9 shows the part of a very dangerous driver's trajectory, where the red line represents the trajectories with violation subtrajectories, the blue line represents the normal trajectory and the red point represents the driver with violation behavior in this position. As mentioned in section 3.3, at least two consecutive points with the violation movement can be characterized with violation behavior for a subtrajectory. So we draw red lines to match the consecutive points with the same violation behavior. It is easy to find that the driver has violation movements on different roads.

The driver's trajectories for a whole day are shown in Fig. 10. It can find that many discontinuous trajectory points have violation behavior and some of points are located at the intersection. Due to the dataset with sampling rate as 30 seconds and having few taxis in the area at the same time, it can infer that the drivers may have illegal behavior such as running a red light. It is not only very dangerous to passengers and themselves, but also to other cars on the road.



Fig. 9. The very dangerous driver.



Fig.10. The very dangerous driver for a whole day.

5 Conclusions and Future Works

Trajectory behavior analysis is becoming very useful in our life. In this paper we presented an algorithm to measure the violation behavior of drivers. Firstly, the algorithm finds violation movements such as speeding and retrograde. Secondly, the driver is divided into different danger levels according to the characteristics values related to the violation movements. Experiments show that this algorithm can correctly detect the violation movement and mark the violation trajectory. Through the experimental analysis, we suggest that the traffic management departments increase supervision at intersections. At the same time, we remind that the drivers do not have violation driving behavior to guarantee the safety of yourself and others.

Tens millions of taxis information are generated every day and it is difficult for us to deal with large amounts of data so we just analysis two days' information. In the future, we will intent to perform more analysis with more real data, to explore speeding, retrograde and other illegal driving behavior such as running a red light.

Acknowledgment

The work was supported by the National Natural Science Foundation of China (No. 61001158, 61272369 and 61370070), Liaoning Provincial Natural Science Foundation of China (Grant No. 2014025003), Scientific Research Fund of Liaoning Provincial Education Department (Grant No. L2012270), Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP no 200801511027), and Fundamental Research Funds for the Central Universities (Grant No. 3132015084).

References

1. WHO - World Health Organization, Global status report on road safety 2015(2015) http://www.who.int/violence_injury_prevention/road_safety_status/2015/zh/
2. http://auto.people.com.cn/n1/2016/0702/c1005-28517746.html
3. Y, Zhu., X, Xiao., J, Zhu.: Research on Active Safety Technology of Automobile Based on Driving Behavior and Intention, in Machinery Design & Manufacture, 2011, PP. 266-268
4. Xu, G., Liu, L., Ou, Y.: Dynamic Modeling of Driver Control Strategy of Lane-Change Behavior and Trajectory Planning for Collision Prediction, IEEE Transactions on Intelligent Transportation Systems, 2012, PP. 1138-1155
5. Kedar-Dongarkar, G., Das, M.: Driver Classification for Optimization of Energy Usage in a Vehicle, Procedia Computer Science, 2012, pp. 388–393
6. Hu, D.I., Zhao, X.H., Mu, Z.C.: Distinguish method of fatigue state based on driving behavior wavelet analysis, Control Conference, IEEE, 2013, pp. 3590-3596
7. Shi, B., Xu, L. Hu, J.: Evaluating Driving Styles by Normalizing Driving Behavior Based on Personalized Driver Modeling, IEEE Transactions on Systems Man & Cybernetics Systems, 2015, pp. 1-1
8. Xu, L., Hu, J., Jiang, H.:Establishing Style-Oriented Driver Models by Imitating Human Driving Behaviors, IEEE Transactions on Intelligent Transportation Systems, 2015, pp. 2522-2530
9. Guo, H., Wang, Z., Yu, B., TripVista: Triple Perspective Visual Trajectory Analytics and its application on microscopic traffic data at a road intersection, IEEE Pacific Visualization Symposium, PACIFICVIS 2011, 2011, pp. 163-170
10. Carboni, E., M, Bogorny,V.: Inferring Drivers Behavior through Trajectory Analysis," Intelligent Systems'2014. Springer International Publishing, 2015, pp. 837-848
11. T, Xu., X, Li.: Analysis on the Safety of Bus Driving Using Vehicle GPS Track Data, GEOMATICS AND INFORMATION SCIENCE OF WUHAN UNIVERS ,2014, pp. 739-744
12. Dueholm, J.V., Kristoffersen, M.S., Satzoda, R,K.: Trajectories and Maneuvers of Surrounding Vehicles with Panoramic Camera Arrays, 2016, PP.1-1
13. He, M., Guo, B., Chen, H.: WhozDriving: Abnormal Driving Trajectory Detection by Studying Multi-faceted Driving Behavior Features, International Conference on Big Data Computing and Communication, 2016, pp.135-144
14. Chen, C., Zhang, D., Castro, P. S.: iBOAT: Isolation-Based Online Anomalous Trajectory Detection, IEEE Transactions on Intelligent Transportation Systems, 2013, pp. 806-818.
15. Krüger, R., Thom, D., Wörner, M.: TrajectoryLenses – A Set-based Filtering and Exploration Technique for Long-term Trajectory Data," Computer Graphics Forum, 2013, pp. 451–460.