Research Article

Grid-Based Anomaly Detection of Freight Vehicle Trajectory considering Local Temporal Window

Zixian Zhang,1,2 Geqi Qi,1,2,3 Avishai (Avi) Ceder,1,4 Wei Guan,1,2 Rongge Guo,1,2 and Zhenlin Wei1,2

1Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China
2School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
3Beijing Research Center of Urban Traffic Information Sensing and Service Technologies, Beijing Jiaotong University, Beijing 100044, China
4Faculty of Civil and Environmental Engineering and the Transportation Research Institute, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel

Correspondence should be addressed to Wei Guan; weig@bjtu.edu.cn

Received 18 April 2021; Revised 23 July 2021; Accepted 12 August 2021; Published 31 August 2021

Academic Editor: Dung-Ying Lin

Copyright © 2021 Zixian Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The security travel of freight vehicles is of high societal concern and is the key issue for urban managers to effectively supervise and assess the possible social security risks. With continuous improvements in motion-based technology, the trajectories of freight vehicles are readily available, whose unusual changes may indicate hidden urban risks. Moreover, the increasing high spatial and temporal resolution of trajectories provides the opportunity for the real-time recognition of the abnormal or risky vehicle motion. However, the existing researches mainly focus on the spatial anomaly detection, and there are few researches on the real-time temporal anomaly detection. In this paper, a grid-based algorithm, which combines the spatial and temporal anomaly detection, is proposed for tracing the risk of urban freight vehicles trajectory by considering local temporal window. The travel time probability distribution of vehicle historical trajectory is analyzed to meet the time complexity requirements of real-time anomaly calculation. The developed methodology is applied to a case study in Beijing to demonstrate its accuracy and effectiveness.

1. Introduction

With respect to urban security, logistics trajectories security plays an important role, which deserves a closer look. In particular, during major events such as the ceremony, concert, and games, real-time monitoring and anomaly detection of the logistics vehicles trajectories have attracted a large amount of security concern since the tragedy in Nice (86 dead and 458 injured) [1] and Barcelona (15 dead and 100 injured). As the vehicle trajectory data can be conveniently collected nowadays by digital sensors, such as various navigation systems, smart cellphones, and RFID devices, researchers can study the reliability of trajectories and the invoked transportation risk issues, i.e., driving behavior detection [2, 3], travel pattern recognition [4–6], and anomaly tracking in practice [4–10].

The unusual change of real-time trajectory may indicate the potential risk in spatial and temporal dimensions. Anomaly trajectory detection aims to find a trajectory that is significantly different from most of the historical trajectories with the same origin and destination. It is mainly reflected in the process of freight vehicles driving from the station to the distribution point. Due to the behavior habits of drivers and the periodicity of urban traffic, freight vehicles have similar driving paths with the same starting and ending points at the same time. In the field of anomaly detection of trajectory, many studies have been investigated from different perspectives [11–13], some of which are used to detect deceptive driving behavior in taxi drivers. However, these studies mainly focus on the spatial anomaly detection, and there are few related to the temporal anomaly detection. In terms of existing research methods, the research on space abnormal
trajectory detection can be segmented into three main categories. The first category is based on distance and density [14, 15]. This category of method usually requires calculating metrics based on the entire historical database, which leads to high time complexity. The second category is based on architecture, mainly neural network architecture and reinforcement learning architecture [16, 17]. This type of method requires a large number of tags, and its accuracy is directly related to the quality of the data. The third category is pattern based [11, 13]. This type of method mainly studies the cooccurrence of trajectory points, so that abnormal trajectories with low probabilities can be detected in a preset window.

However, there are few related researches on real-time temporal anomaly discovery, which can help find the abnormal location of vehicles from the temporal perspective. This may be due to the following reasons: (1) urban transportation network flow is dynamic and complex, which leads to the different temporal security assessment threshold corresponding to different spatial locations, such that intersection parking and road section parking are different; (2) spatial anomaly can be judged by the isolated trajectory point, while temporal anomaly detection required a detection window cooperating multiple sequential trajectory point; (3) in real-time detection, selecting an appropriate detection window size is critical, and unreasonable selection may lead to completely wrong judgment. In extreme example, with only one trajectory point in the window, the travel time anomaly of a single trajectory point may be slight. However, with the expansion of time window, it is possible to exceed the threshold set by the detection window of historical trajectories.

In light of this, the window time of isolation based online anomalous trajectory detection (WT-iBOAT) algorithm is proposed in this paper to accomplish the real-time anomaly detection of abnormal logistics trajectories. The research can help supervise the driver of logistics vehicles to drive the vehicle safely and reliably and enable the real-time anomaly detection of freight vehicle trajectories by considering both spatial and temporal variations. The main contributions are as follows:

(i) First, we extract different constraints from not only spatial, but also temporal dimensions for comprehensively solving the motion-based anomaly detection problem

(ii) Second, a grid-based anomaly detection method is proposed from a new perspective of using the local temporal window, which is beneficial for increasing calculation accuracy and reducing the computation complexity of real-time detection

(iii) Third, this paper proposes a framework for detecting real-time risk trajectories of urban freight vehicles, and a real-world data set is utilized to verify its feasibility

The paper is organized as follows. In Section 2, the relevant work is investigated. In Section 3, related architectures and algorithms are proposed. In Section 4, the results of three trajectory datasets are analyzed using the algorithm proposed in this paper. In Section 5, the performance of the algorithm is analyzed. In Section 6, the the adjustment mechanism of the algorithm is discussed. In Section 7, we summarize and discuss the whole paper.

2. Related Research

The trajectory data have been widely used in the popular route discovery and anomaly detection. In the field of popular route discovery, Zaibin et al. [18] proposed a coherence expanding algorithm and a maximum probability product algorithm to discover the most popular route. Jani et al. [19] used a fast map server, which could generate and visualize hot maps of hot routes online from massive motion trajectory data according to customers’ preferences. Luo et al. [20] found that suffix-optimal, length-insensitive, and bottleneck-free methods reflect common-sense notions and proposed efficient search algorithms to determine the most frequently used paths based on the time period. The semantic features mainly come from other moving objects with similar behaviors. Treerapot et al. [21] proposed a statistical approach to estimate truck activities and freight analytics from Global Positioning System (GPS) data of trucks and applied the method in estimating activities, activity patterns, commodity trip chains, and status of trips legs from the collected truck GPS data. Wei et al. [22] studied the characteristics of uncertain trajectories and proposed a route inference framework based on a collective knowledge algorithm to find the top-k popular routes. Outlier detection is important work in mining trajectories [23, 24]. Lee and Cho et al. [24] proposed a grid-based local outlier factor algorithm for outlier detection. This algorithm reduces the computation time significantly with acceptable trade-off errors. Knorr et al. [23] studied the notion of distance-based outliers to find outliers in multidimensional datasets. Discovering the popular route is always used for daily management and operation of urban traffics. Sometimes, it is helpful to find out the abnormal trajectories, but it lacks directness compared with anomaly detection and may result in misidentification of some regular trajectories (not most popular ones).

In the field of anomaly detection, Wang et al. [25] used improved Robust Principal Component Analysis (RPCA) to introduce the detection of abnormal passenger flow and Spatial Temporal-Density-Based Spatial Clustering of Application with Noise (ST-DBSCAN) algorithm to group the detected station-level anomalies. Suzuki et al. [26] applied the hidden Markov model (HMM) and k-means, and Bera et al. [7] used Bayesian learning techniques to detect abnormal pedestrian trajectories. Other popular fields related to GPS traces include the detection of social communities [27] and abnormal driver behavior [28]. Yin et al. [27] proposed a probabilistic model, which was used to infer users’ social communities by integrating their spatiotemporal data and semantic information. Ali and Ismail [28] investigated driver behavior through GPS trajectory data from the vehicle, mainly considering the features of speed limit, sudden speed changes, and constant changes in the
vehicle lateral position, and they used these features to detect abnormal driver behavior.

Many researchers have conducted related research on abnormal trajectory detection. In Song et al. [16], in order to solve the problems, i.e., the traditional anomaly trajectory detection method cannot adequately capture the sequence information of the trajectory, and focusing only on the given source and target may lead to data sparseness. They proposed an algorithm that used an Anomalous Trajectory Detection using Recurrent Neural Network (ATD-RNN) to represent the trajectory by learning a trajectory embedding and then found the risk trajectory. Leping et al. [15] put forward a trajectory segmentation strategy based on multimotion features and a similarity measurement method based on trajectory structures. With the proposed strategy and method, a distributed trajectory clustering algorithm is designed to improve the efficiency of the clustering algorithm. To tackle the problem of calculating distance and neighborhood density in the trajectory anomaly detection algorithm, the dataset was pruned by the trajectory clustering results, which improves the efficiency of the algorithm. Nonetheless, Song et al. [16] and Leping et al. [15] required artificial label of abnormal trajectory for improving effective of algorithm and could not perform real-time anomaly detection.

Wu et al. [17] proposed a probabilistic model-based vehicle outlier detection method, Driving Behavior-based Trajectory Outlier Detection (DB-TOD), which was used to model the driving behavior in a historical trajectory set. They designed outlier detection algorithms for both full and partial outliers, which solved the problem for providing early warnings of outliers only when the outliers were partially observed. Lee et al. [8] et al. aimed to address the problem that the detection algorithm for tracking outliers has limited ability in detecting the outliers’ tracks. An outlier detection algorithm, time-dependent popular route-based trajectory outlier detection (TPRO), was proposed. The algorithm was divided into two stages: partitioning and detection. For the first stage, a two-layer trajectory partition strategy has ensured the high quality and high efficiency of subtrajectory partitioning. In the second stage, the hybrid outlier detection method based on distance and density clustering improves the accuracy of the trajectory outlier detection algorithm. Daqing et al. [11] proposed an isolation-based anomalous trajectory (iBAT) method based on isolation. First, the original trajectory was mapped into the grid sequence, and the historical grid trajectory index was established. Then, the iForest algorithm was used to detect the isolated trajectory. However, this approach can only support full trajectories. Sun et al. [29] proposed the iBOAT algorithm based on trajectories, similar to the real-time detection method, which was an influential method. This method was developed from the iBAT algorithm. The main improvement of this method is the way that it compares with iBAT in view of the whole line offline, as it is improved by using the method of real-time contrast. It is more practical, because it can be a fast detection algorithm with high reliability. Yu et al. [30] defined anomalous behavior of moving objects according to the idea of neighbor-based trajectory outliers and proposed the minimal examination framework to effectively discover new abnormalities. Yingyi et al. [31] proposed a method of monitoring distance-based anomalies over trajectory streams. This method mainly used local clustering and piecewise vantage point trees with the corresponding pruning method to reduce the calculation time. These studies can make efficient detecting trajectory outlier while preserving the effectiveness. But these algorithms only study space anomaly detection and do not relate the temporal anomaly detection.

Jie et al. [12] used a historical trajectory dataset and commonly detected abnormal trajectories. Then, they modified the TPRO algorithm by dividing the steps into offline processing and online detection and adding appropriate modules to adapt the detection. This paper proposed an algorithm named trajectory outlier detection (TPRRO). TPRRO is a real-time outlier detection algorithm, and it constructs short for time-dependent transfer index (TTI) and short for time-dependent transfer graph (TTG) to conduct spatiotemporal anomaly detection. Although this paper proposed an algorithm for temporal anomaly detection, it did not analyze temporal anomaly detection in detail. At the same time, Jie et al. [12] used time-dependent transfer graph (TTG), which calculated the travel time from a vertex to another in road network, so this algorithm could not reflect immediate temporal anomaly information of trajectory.

In summary, we present a real-time temporal anomaly detection algorithm for logistic vehicle to prevent potential security incidents. Different from the previous study (in Table 1, we summarize contributions of previous study and our research), we focus on the real-time temporal anomaly detection method and design algorithmic framework for spatiotemporal anomaly detection of Urban Freight Vehicles. We further conduct analysis to our collecting data and explain the possible motivations and reasons of the anomalous behaviors.

3. Methods

In this section, the algorithmic framework of this paper is introduced, including data preprocessing (DP), offline processing (OFP), the data manager (DM), and online processing (OLP). The purpose of DP is to obtain the effective loading and unloading positions of raw GPS data. The goal of OFP is to obtain the available subtrajectories from the loading (origin) and unloading (destination) positions obtained from the DP and to mapping the subtrajectories to a preset grid. The goal of OLP is to real-timely meshing test trajectory points. The DM is mainly used for extraction of spatiotemporal experience constrained from the historical trajectories and detect abnormal trajectories. Figure 1 shows the submodules within the modules.

3.1. Concept Relationship. This section first introduces the relationship among the concepts involved. For trajectory anomaly detection, the relationship between trajectory point and time/loading position/unloading positions is many-to-one, because many trajectory points are associated with the
same time/loading position/unloading position. For the
same loading and unloading positions, the trajectories are
different in different time periods. A loading position may be
related to the trajectory of different roads, different times,
and different unloading positions. In addition, trajectories
crossing a specific road section may involve different loading
positions; trajectories crossing a specific road section and
different unloading positions may also involve similar
relationships.

3.2. Meshing of the Spatiotemporal Trajectory Point. The
movement of vehicles produces a lot of trajectory data. We
need to divide the existing historical trajectory effectively
before detecting the spatiotemporal anomaly of trajectory.
The granularity of spatiotemporal trajectories is crucial when
dividing trajectories. Different mesh sizes are suitable for
different trajectory granularity problems. This is the reason
why we use the grid method, which does not require us to
know all the space-time coordinates, which avoid a lot of
calculation. The trajectory point is mapped to the two-di-
mensional space of the grid cell (Figure 2). This is a two-
dimensional Euclidean space defined by x-axis and y-axis. The
center of each grid cell is used to represent all GPS points of
the grid cell. If the grid cell is too large, the accuracy of
anomaly detection may not be enough. However, the smaller
grid cell may lead to much computation. Therefore, we must
find an optimal grid size that can identify trajectories without
unnecessary computation. Based on two-dimensional grid
division, the trajectory query from specific loading position
coordinates to specific unloading position coordinates be-
comes the trajectory query from one grid cell to another.

We define the minimum range of grid as (COLNmin, ROWNmin, COLNmax, ROWNmax) and the length range
of a grid as k. X and Y present the numbers of columns and
rows of grid in the study area [32].

\[
X = \frac{\text{COLN} \max - \text{COLN} \min}{k},
\]

\[
Y = \frac{\text{ROWN} \max - \text{ROWN} \min}{k}. \tag{1}
\]

3.3. Rough Selection of the Trajectory Set. Before anomaly
detection of test trajectory, we first need to filter the historical
trajectory, which has the same loading and unloading point as
the test trajectory. We define this process as the query of
historical trajectory. The related concepts are defined as follows:

**Definition 1** (trajectory). Trajectories \( \text{tr} = [P_1, P_2, ..., P_N] \) are
collected GPS records. Each GPS record \( P_i \) specifies the
instantaneous time, position (latitude and longitude), and
vehicle ID of the vehicle.

**Definition 2** (subtrajectory). The subtrajectory reflects dif-
ferent driving preferences of different drivers for vehicles
with the same transportation purpose in different periods.
We use the effective parking time as the division factor of the
subtrajectory. So, the subtrajectory is a part of the trajectory.

**Definition 3** (scalable loading and unloading point of tra-
jectories). In order to eliminate irrelevant trajectories, the
loading and unloading points are used as spatial constraints
to filter all trajectories. In order to achieve this goal, we first
find loading points for all the subtrajectories and keep the
subtrajectories. After that, we find the corresponding
unloading point in kept subtrajectories. Compared with the
common input method, this method has a huge advantage: a
trajectory is used in anomaly detection when the middle
section of this trajectory has the same loading and unloading
points as the test trajectory, as shown in Figure 2.
3.4. Extraction of Spatiotemporal Constraint

**Definition 4** (global spatial constraint). The different drivers have specific road choice preference in a certain period; this difference comes from drivers’ perception of the road traffic system. However, there are some uncertainties in this perception. For experienced drivers, the optional travel in a fixed time is regular and searchable. Specifically, the travel with the same loading and unloading positions generally follows optimal driving conditions. For example, the driver may require the minimum distance, time, and fuel consumption of the travel path. The example can be seen in Figure 3: O is the loading position, and D is the unloading position. The purple line represents the historical trajectory, and the red dotted line represents the test trajectory. The blue box represents the grid of online mapping. The yellow box indicates the coordinates of the grid.

In order to extract the spatial constraints of trajectory, we use part ideas of IBOAT algorithm. The specific methods are as follows. Suppose that there are three purple trajectories in Figure 3, \( T = \{(1, 1), (5, 2), (7, 2), (10, 3), (13, 3), (17, 3), (21, 3), (24, 3), (27, 4)\}, \( t_2 = \{(1, 1), (5, 2), (8, 2), (11, 3), (15, 3), (19, 3), (23, 3), (26, 3), (27, 4)\}, \) \( t_3 = \{(1, 1), (5, 2), (7, 2), (11, 3), (13, 3), (19, 3), (21, 3), (25, 3), (27, 4)\}, \) and a test trajectory \( t = \{(1, 1), (5, 2), (8, 2), (11, 3), (19, 3), (27, 3)\}, \) then hasPath \((T, t) = \{t_2, t_3\}. \) So, hasPath is defined as a function that returns the set of trajectories from \( T \) (historical trajectory) that contains all of the points in \( t \) (test trajectory) in the correct order [29].

\[
\text{hasPath}(T, t) = \left\{ t' \in T \left| \begin{array}{l}
(1) \ 1 < i < n, \ t_i \in t' \\
(2) \ 1 < j < n, \ t'_j < t'_j
\end{array} \right. \right\}
\]

The support of the test trajectory is calculated as follows: in the grid coordinates (3, 2), the red trajectory and purple trajectory are in the same grid, and the support value of the test trajectory is 100%. The grid coordinates (9, 5) contain only the red trajectory. The support value of the test trajectory is 0% [29].
The support has a great influence not only on the spatial anomaly detection of trajectory, but also on the temporal anomaly detection. The modeling of real-time temporal anomaly depends on the spatial anomaly detection results. In other words, temporal anomaly detection is performed when the space is normal. Therefore, this condition is a strong constraint for temporally anomaly detection.

**Definition 5** (local spatio constraint). The support considers the global spatial characteristics of trajectory but lacks the local spatial characteristics. The trajectory is composed of trajectory points that have different effects on the abnormality of the current trajectory point. In short, the closer the trajectory point is to the current test trajectory point, the greater the impact on the anomaly. On the contrary, the smaller the impact. In view of this characteristic, this paper adds two parameters, namely, \( t_2 \) (length of the maximum continuous temporal anomaly...
window) and \( t_1 \) (length of the warning time window), to correct the results of the spatial anomaly. (In Section 6.2, we will explain in more detail.)

\[
R_t = 100 \times \frac{\min(N_n(t_1), N(t_2))}{N(t_2)}. \tag{4}
\]

Equation (4) is the model for calculating \( R_t \) (the risk value of the trajectory). \( N_n(t_1) \) is the maximum number of abnormal GPS points of the test trajectory in \( t_1 \). \( N(t_2) \) is the number of all GPS points of the test trajectory in \( t_2 \). The longer the \( t_1 \) is, the lower the tolerance of the model, while the shorter the \( t_2 \) is, the higher the sensitivity.

**Definition 6** (global temporal xonstraint). Since the road network and traffic conditions are dynamic, the trajectory is also dynamic and depends on the traffic conditions. Therefore, the trajectories of the same loading and unloading positions may be completely different during peak hours and other times of the day. In terms of time, we divide trajectories according to the traffic conditions in different periods. Taking Beijing as an example, the morning and evening peaks are from 7:00 a.m. to 9:00 p.m. and from 5:00 p.m. to 8:00 p.m. Based on this, we divide the time into four periods: 7:00 am–9:00 am, 9:00 am–5:00 pm, 5:00 pm–8:00 pm, and 8:00 pm–7:00 am. We look for the same time as the test trajectory.

**Definition 7** (local temporal window constraint). The global temporally constraint classifies the trajectories from the overall time, but this is not sufficient for the real-time anomaly detection of the trajectory. Since the global temporally constraint span is too large to meet the real-time accuracy requirements of the trajectory, the extraction of the local temporally constraint with a smaller span will be the key to real-time anomaly detection.

A window time is used in equation (4), which inspired the idea of using sliding windows for temporal anomaly detection. The different window sizes provide different information, which is beneficial in the judgment of temporal anomalies. There are two types of sliding windows. One has a fixed time size, but the number of grids in the window is variable; the other has a fixed number of grids in the window, but the time is variable. This paper uses the second method, which is called fixed position window, because it is more intuitive to compare the travel time of the historical trajectory and the test trajectory in the same window.

**Definition 8** (Construct Empirical Local Temporal Window Constraints Based on Historical Information). According to the travel time preference obtained from all historical trajectories, travel temporal is divided into normal temporal and abnormal temporal. Based on the method of detecting outliers using box plots in anomaly detection, inspired by this, we use quantiles as a statistical reasoning tool to determine the abnormalities of travel temporal in window. In the historical trajectory, the trajectories that meet the conditions of Definitions 4 and 6 constitute the travel temporal frequency of all historical trajectories under a preset size window.

Let function \( c(t) \) be the travel temporal frequency for window, determined by counting the length of travel temporal in window that meets the above conditions in all trajectories.

\[
p(r) = P(c(t) = r), (r > 0). \tag{5}
\]

Let \( M(r) \) be the number of windows that satisfy \( c(t) = r \). Let \( Y \) be the total number of windows. Then, the formula can be rewritten as

\[
p(r) = \frac{M(r)}{Y}. \tag{6}
\]

Let \( f(q) \) be the cumulative distribution function of \( r \).

\[
f(q) = \int p(r)dr, (r \geq 0), \tag{7}
\]

Let \( Q_k(T_i, t_i) \) be the k-th quantile of the travel time probability distribution of the historical trajectories in the \( i \)-th window. We use a probabilistic method to model the distribution of the travel times of historical trajectories to detect temporal anomalies in the test trajectory.

\[
HT_i = \frac{\text{hasPath}(T_i, t_i)}{|T_i|} > \theta. \tag{8}
\]

**Definition 9** (adjustment mechanism: local temporal window constraint extraction based on redundant information). In the anomaly detection problem, real data may occur, in which historical trajectory (HT) travel time is greater or less than the new test trajectory (NTW) travel time in the \( i \)-th window. In this case, the travel time in the \( (i-1) \)-th window impacts the travel time in the \( i \)-th window. Therefore, we want to eliminate this error. We define this case as follows: the test trajectory in the window has redundant information. This effect is significant for temporal anomaly detection, which will lead to insufficient or overloaded historical information in window. These two situations directly lead to the failure of constraint extraction or not. Because the travel time of the \( (i-1) \)-th window affects the travel time of the \( i \)-th window, therefore, we want to eliminate this error by using a linear equation and Definition 8.

Specifically, when HT is greater than NTW, if the HT value in the \( i \)-th window is not adjusted, the information in the \( (i-1) \)-th window may be carried into the \( i \)-th window. This causes the HT value in the \( i \)-th window to be greater than the required HT. An adjustment coefficient \( (r_p) \) is needed in this situation.

HT is less than NTW in Figure 4. This is the opposite of the above situation, but it will also cause errors due to the introduction of redundant information. Therefore, the same needs to be adjusted.

The formula is as follows:

\[
HT_i = HT_i + r_p \sum_{j=1}^{su} \Delta t_{pj} + r_n \sum_{j=1}^{nu} \Delta t_{nj}. \tag{9}
\]

The difference between the travel time of the historical trajectory and the travel time of the test trajectory is...
Finally, we obtain the rule that the locations where \( t_7 \) is followed by \( t_5 \) in TW8 in the historical trajectory, there is a 30 second difference between \( t_4 \) and \( t_7 \). \( t_4 \) means that time steps (5, 6, 7) are in the same grid. \( k \% \) here is the value of HT in NTW. However, by comparing NTW and HT, the change in the travel time of the test trajectory and the historical trajectory follows:

\[
\Delta t_i = NTW_i - HT_i. \tag{10}
\]

In equation (10), \( HT_i \) indicates the value of \( k \)-th quantile for travel time in \( i \)-th window.

Trick (relaxation for HT less than NTW): The case where HT is greater than or less than NTW has different effects on abnormal judgment. More precisely, when HT is greater than NTW, it means that the test trajectory is normal, but the redundant value will affect the abnormal judgment of the next point. When HT is less than NTW, the test trajectory is abnormal. This is a situation where we need to detect time anomalies. At this time, the estimation accuracy of the redundant value is fatal to the result of abnormal judgment, so we try to relax the constraint at this time. Figure 4 shows this trick method. An example is given to illustrate the above method. Figure 4 shows the calculation results of this method when the window size is 3.

In Figure 4, HT represents the travel time of the historical trajectory. The digital value of HT represents the time step—a time step is 10 seconds—and the grid represents the location of the vehicle. NTW is the travel time of the test trajectory in the window. TW is the \( i \)-th window; \( \Delta t_i \) is the difference between NTW and HT, which represents the difference between the travel time of the test trajectory and the historical trajectory. \( \Delta t_i \) is the difference between \( \Delta t_i \) and \( \Delta t_{i-1} \). In Figure 4, TW8 is a special window. Time Step 4 is followed by time Step 7 in TW8 in the test trajectory, which means that time steps (5, 6, 7) are in the same grid. There is 30 seconds between time Step 4 and time Step 7. Time Step 4 is followed by time Step 5 in TW8 in the historical trajectory, which means that only 10 seconds elapses. If NTW and HT are directly compared at TW8, this will lead to an incorrect conclusion. Through observation, we discover that this incorrect conclusion is caused by time steps 5, 6, and 7 in TW8 in NTW. However, by comparing \( \Delta t_i \), this problem can be solved. The value of \( \Delta t_6 \) is 0, and the value of \( \Delta t_7 \) is (–20). Finally, we obtain the rule that the locations where \( \Delta t_i \) is greater than 0 are abnormal, which are marked in red; the locations where \( \Delta t_i \) is less than or equal to 0 are normal and are marked in blue. Black indicates a normal value.

**Definition 10** (constructing empirical average speed ratio constraint). The change in average speed can express the change in travel time well. Therefore, the speed is referenced to help filter out the temporal anomaly in equation (6), where \( a \) is the speed ratio threshold, and WTR is the speed ratio of the \( i \)-th window. When WTR is greater than \( a \), the travel time of the \( i \)-th window is abnormal. The formula is as follows:

\[
\text{WTR}_i = \left| \frac{\text{speed}(Q_i([\text{hasPath}(T_i,t_i)])/|T_i| < \theta)) - \text{speed}(NTW_i)}{\text{speed}(Q_i([\text{hasPath}(T_i,t_i)])/|T_i| < \theta))} \right| > a. \tag{11}
\]

**Definition 11** (travel time of the test trajectory). NTW is the travel time of the test trajectory, which is equal to the length of the window minus 1.

\[
NTW_i = (|\text{len}(TW_i)| - 1). \tag{12}
\]

The above description is called window time of isolation based online anomalous trajectory detection (WT-iBOAT) algorithm, which adds an abnormal temporal detection method to iBOAT [29]. Algorithm 1.

The pseudocode of WT-iBOAT algorithm is as follows:

**4. Results**

**4.1. Data Description.** To detect abnormal trajectories of urban freight vehicles, the historical trajectories provided by a logistics company in Beijing were used in this paper.

The data were collected from approximately 35 vehicles in Beijing in 2019 (September to October 2019). The main fields are the vehicle status, longitude, latitude, offset
dimension, city code, working condition time, GPS time, accumulated mileage, positioning status, vehicle speed, vehicle ID, etc. It should be noted that these data are recorded every 10 seconds. We first need to segment the raw trajectory according to the business requirements of logistics. Therefore, loading and unloading points that are common locations for vehicles to load and unload goods are used as the segmentation points of the trajectory. Loading and unloading points are obtained by the effective parking time of the vehicle. The effective parking time depends on the experience of the staff. The regularity of vehicle trajectory can be found by loading and unloading position obtained from the effective parking time. Clustering is a good choice to obtain this regularity. DBSCAN [33] is a density based algorithm, which assumes that the category can be determined by the compactness of the sample point distance.

DBSCAN is commonly used to find the loading and unloading positions. The parameters of the algorithm are minimum number of fields (MINPTS) = 60 and domain radius (EPS) = 150 meters. A total of 151 loading position categories and 151 unloading position categories are obtained. The clustering results based on the ArcGIS platform are shown in Figure 5. On the left is the loading position, and on the right is the unloading position. The loading and unloading positions in the central city are relatively concentrated, which meets the requirements of urban logistics, indicating that the clustering results reflect the common working area.

By screening the above loading and unloading positions, three trajectory datasets with relatively good quality are finally selected to test the algorithm. Figure 6 describes the data, including the visualization of the trajectory data, the travel time distribution, and the travel distance distribution.

To analyze the performance of the spatial anomaly detection of the algorithm, we manually mark all trajectories and then calculate the accuracy of the algorithm. The performance analysis of temporal anomaly detection uses the method of comparing the different window sizes.

Figures 7 and 8 show visualizations of spatiotemporal anomalies using a hot map. The hot map, which uses different colors to indicate statistically abnormal values, is generated by counting the number of anomalous grids in the window on a test trajectory. The red points are GPS points, which represent the travel route of the vehicle. Figure 8 displays only the nonabnormal routes.

### 4.2. Results of Spatial Anomaly Detection

In Figure 7, different spatial anomalies are highlighted. The normal trajectories are almost the shortest routes between the loading and unloading positions. An analysis of each picture is given below. In Figure 7(a), a route with more serious spatial anomalies chooses the East 4th Ring Expressway instead of the East 5th Ring Expressway. There are 106 trajectories in this dataset, 6 of which are on this route. Through map matching, it is found that these trajectories do not leave the East 4th Ring Expressway. Therefore, these abnormalities may be caused by the driver’s negligence, which leads to a failure to change lanes in time and a missed opportunity to enter the 5th Ring Expressway. In Figure 7(b), there are some locations with serious abnormalities and concentrated areas. This may be due to vehicles with temporary

**Algorithm 1: WT-iBOAT.**

```plaintext
Input: input point index, WL: window time step, T: historical trajectory dataset, θ: spatial anomaly threshold, α: temporal anomaly threshold
Output: X: abnormal trajectory points, R_i,: the risk value of the trajectory

Process:
(1) While new grid received from online gridding
(2) support (i) = hasPath (ti−1, ti) / T
(3) if support (i) < θ:
(4) X ← update (support (i))
(5) T ← update (hasPath (ti−1, ti))
(6) else:
(7) HT ← Q_i (hasPath (T_i, ti)) / |T| < θ
(8) HT ← update (HT, |T|, Δt, r)
(9) NTW ← calculate (cell_i, cell_j, α)
(10) Δt ← calculate (NTW, HT)
(11) Δt ← update (Δt, Δt)
(12) if WTR > α and Δt > 0:
(13) X ← update (WTR, Δt)
(14) end if
(15) end if
(16) R_i ← calculate (count (X), count (WL))
(17) end while
```
Figure 5: Clustering results of the loading and unloading points (map comes from https://www.tianditu.gov.cn/).

Figure 6: Data description (map comes from https://www.tianditu.gov.cn/). (a) Visualization. (b) Travel time distribution. (c) Travel distance distribution.
transportation tasks. In Figure 7(c), the purpose of detouring vehicles may be to avoid large intersections. However, Figure 8(c) shows that there are not many temporal anomalies at the intersection. Instead, there is a temporal anomaly on the street near the destination. Therefore, it is reasonable that the vehicle is avoiding possible traffic jams on that street.

4.3. Results of Temporal Anomaly Detection. In Figure 8, red represents abnormally severe areas in terms of anomalies, and blue represents abnormally mild areas. From Figure 8, it can be concluded that most of the abnormal areas are close to the intersection, and the areas close to the starting and ending points are more likely to be abnormal.

In Figure 8(a), the temporal anomalies mainly occur at the intersection of Beijing Shenyang Road and Huosha Road and at the intersection of Beijing Shenyang Road and the airport expressway. A road with almost no temporal anomalies is the East 5th Ring Expressway. The temporal anomalies in Figure 8(b) mainly occur in the suburbs between Hebei Province and Beijing city and at the intersection of Guangming West Road and Xichang Road in Langfang city. In the suburbs between Hebei Province and Beijing, the roads pass through settlements, and temporal anomalies occur in dense residential areas. The intersection of Guangming West Road and Xichang Road in Langfang city has a large traffic volume and is prone to traffic congestion. Figure 8(c) shows the Shangqing Bridge in Beijing city. The figure shows that temporal anomalies mostly occur on the roads approaching adjacent residential areas and at interchanges (such as the intersection with Shangqing Bridge), which is consistent with the actual situation.

The above analysis describes the reasons for spatio-temporal anomalies under different conditions. It is demonstrated that the anomaly detection algorithm proposed in this paper can address actual situations.

4.4. Results of Abnormal Trajectory Detection. Finally, a trajectory is used for qualitative analysis to present all the results of the algorithm. Figure 9 shows a trajectory with ID 27636. The spatial shape of this trajectory is shown in Figure 9(a). The blue points are normal, and the red points are abnormal.

Figures 9(b)–9(d) display the temporal anomalies, spatial anomalies, and risk value of the trajectory, respectively. Figure 9(d) expresses all the spatiotemporal anomaly information. A change in color indicates a change in the density of the temporal (Figure 9(b)) and spatial (Figure 9(c)) anomalies in $t_i$ of equation (1). From Figure 9(d), we can identify anomalous trajectory segments.
and determine a level of risk to warn of the potential security risk. Figure 9(d) is the concept of risk value calculated according to equation (4). Users can define the risk level according to their own needs according to different value of risk. In equation (4), $t_1$ and $t_2$ can be adjusted according to the varied demand of different scenarios.

5. Analysis

This section comprehensively analyzes the impact of different parameters on accuracy and analyzes the differences between our proposed method and iBOAT and TRAOD algorithms. After that, the results of the three methods are explained on our data.

Environment: our algorithm is mainly implemented in Python. The machine used for the experiment has a 9th generation i7 CPU and 32 G memory. The operating system is Windows 10 x86_64, and the compiler is Spyder 4.15.

5.1. Spatial Anomaly Parameter Analysis. Spatial anomaly detection for vehicles focuses mainly on the impact of the grid size on accuracy and processing time. Figure 10 displays these results. In Figure 10, the blue line represents the impact of the grid size on the accuracy, and the red line shows the impact of the grid size on the processing time of the program. The x-axis is the grid size. The left y-axis is the number of normal trajectories. The right y-axis is the processing time. A representative dataset is selected considering the long transport distance of urban freight vehicles. There are 106 trajectories in this dataset, of which 99 are normal and 7 are abnormal. Figure 10 shows that when using a 250 m × 250 m grid, most normal trajectories were found, and the processing time was at a high level. Therefore, 250 m is selected for spatial anomaly detection.

5.2. Temporal Anomaly Parameter Analysis. For temporal anomaly detection, manual methods are not used, because it is difficult for manual methods to find a reasonable standard to judge when and where the GPS points are abnormal. Therefore, we observe the impact for other window sizes, with the result of window size = 2 as the benchmark, and we conduct a contrastive analysis to explain the feasibility and advantages of the algorithm proposed in this paper. Figure 11(a) shows the method, where the x-axis represents time and the y-axis represents space. Spaces 5 to 3 (the window is 2) are used as a benchmark to judge the algorithm’s performance under different parameters, e.g., spaces 5 to 1 (the window is 4). Next, we will analyze the performance and accuracy of WT-iBOAT.

Performance: Figure 11(b) shows the processing time of WT-iBOAT and iBOAT. The processing time of WT-iBOAT is approximately 2.5 times that of iBOAT. It does not change
with the number of historical trajectories. This shows that the processing time is stable, and the processing time of each GPS point is approximately 0.006 seconds. Therefore, the algorithm can be used for real-time detection.

Accuracy: the travel time of a trajectory will fall into one of four cases: true positives (TP) are the number of anomalous travel times of trajectories that are classified as anomalous; false positives (FP) are the number of normal travel times of trajectories that are classified as anomalous; true negatives (TN) are the number of normal travel times of trajectories that are classified as normal; and false negatives (FN) are the number of normal travel times of trajectories that are classified as anomalous. We report these cases and calculate the common metrics of precision, true positive rate (TPR), defined as TPR = TP/(TP + FN), and true negative rate (TNR), defined as TNR = (TN/TN + FP). A perfect classifier will have TNR = 1 and TPR = 1.

In Figure 12(a), window size ≤ 6, quantile = 75, TPR > 0.75; in Figure 12(b), quantile > 75, window size = 4, and TPR > 0.825. This reflects that as long as the parameter settings are reasonable, our algorithm can achieve relatively excellent results. To be more specific, when considering the effect of the window size on accuracy, the quantile is determined to be 75. This is because when the quantile is 75, there is a higher requirement for the travel time, which is convenient for comparing the results of the algorithms. The 75th quantile and the 2 windows are selected as the target values. Figure 12(a) plots the results with regard to different values of window size. TPR decreases rapidly with increasing window size, which indicates that the larger the window size is, the longer the travel time allowed by the window, so the proportion of positive sample detections decreases. Quantifying TPR, when the window sizes are 3, 4, 5, 6, 7, 8, 9, and 10, the TPR calculated by the algorithm is 87%, 82.5%, 78%,
77%, 75%, 74%, 72%, and 71% of the benchmark, respectively.

However, TNR hardly declines, indicating that the window has little impact on the judgment of negative samples, which is actually the role of the adjustment mechanism in the algorithm. The adjustment mechanism effectively reduces the impact of previous moments in the i-th window.

When studying the impact of the quantiles on accuracy, the window size is chosen as 4, because this is an obvious turning point in Figure 12(a). This shows that enough information is considered in 4 windows. Figure 12(b) shows that TPR and TNR increase with increasing quantiles. This means that the higher the tolerance is, the less temporal anomaly detection occurs. Specifically, the TPR of 4 windows and 70th quantile is 82% of that of 2 windows and 70th quantile; the TPR of 4 windows and 75th quantile is 82% of that of 2 windows and 75th quantile; the TPR of 4 windows and 80th quantile is 85% of that of 2 windows and 80th quantile; the TPR of 4 windows and 85th quantile is 87.5% of that of 2 windows and 85th quantile; the TPR of 4 windows and 90th quantile is 91.5% of that of 2 windows and 90th quantile; the TPR of 4 windows and 95th quantile is 95% of that of 2 windows and 95th quantile.

5.3. Abnormal Trajectory Detection and Analysis. In order to illustrate the algorithm more effectively, we compare our
algorithm with the TRAOD algorithm and iBOAT algorithm in the case study. TRAOD is an offline algorithm, which can get abnormal trajectories by clustering the segments of different trajectories [8]. The TRAOD algorithm has high efficiency on finding out the trajectories with more lane changes. The results are as follows: in 212 trajectories, our method found all 48 abnormal trajectories, while iBOAT algorithm only found 13 abnormal trajectories, and TRAOD algorithm found 39 abnormal trajectories. Among them, 11 of the 48 abnormal trajectories found by our method are the same as 11 of the 39 abnormal trajectories found by TRAOD, of which 4 belong to spatial anomaly and 7 belong to temporal anomaly. iBOAT and TRAOD have the same four spatial anomaly trajectories. All the 13 abnormal trajectories found by iBOAT algorithm are in 48 trajectories found by our method. The results are shown in Table 2.

6. Discussion

In this section, we discuss the role of the adjustment mechanism in the algorithm. Then, an example is used to explain how equation (1) improves the results of spatial anomaly detection.

6.1. Overall Discussion of the Adjustment Mechanism.

For the overall discussion of the adjustment mechanism, the speed ratio can explain the information of the window well. If speed ratio > 0, the travel time is abnormal. If speed ratio < 0, the travel time is normal. The larger the speed ratio is, the more serious the temporal anomaly, and the smaller the speed ratio is, the milder the temporal anomaly. The following case needs to be explained: when the test trajectory immediately turns a spatial anomaly into a temporal anomaly, the travel time of the historical trajectory is zero, and the speed ratio is infinite. To detect this, when calculating the speed ratio, greater sensitivity is required. Therefore, the historical speed is taken as 50 m/s, so the first speed ratio is equal to (50-250)/10= -5, and the subsequent speed ratio will also be greater than 0.5.

Figure 13 is the kernel density diagram of the abnormal speed ratio of all datasets. The green, yellow, black, and blue areas are the kernel density diagrams of the 75th quantile with a window of 4, the 90th quantile with a window of 4, the 75th quantile with a window of 8, and the 90th quantile with a window of 8, respectively. The first row is the kernel density diagram without the adjustment mechanism, and the second row is the kernel density diagram with the adjustment mechanism.

Note that the different color areas in the first row are more dissimilar than their counterparts in the second row. This shows that the adjustment mechanism makes the abnormal speed ratio more reasonable in different situations. The adjusted results tend to be more concentrated in areas with smaller speed ratios, and the larger areas are reduced accordingly, which is a manifestation of the rationality of the adjustment.

6.2. Improved Analysis of Spatial Anomaly.

The algorithm judges whether the test trajectory is a spatial anomaly by calculating the support values. However, simply using the support values causes errors due to the grid. Figure 14 shows the results of the problem and improvement. The different colors in the left figure represent different support values. Red represents larger support values, and blue represents smaller support values. In the green box, when the grid size is 250 × 250 meters, a normal spatial point may be judged as an abnormal point. Section 4 equation (1) can be used to solve this problem. The right figure shows the result of using this formula when \( t_2 \) is 10 minutes and \( t_1 \) is 3 minutes. The right figure reflects the correction of the calculation results very well. Blue indicates a low abnormality, and red indicates a high abnormality.

<table>
<thead>
<tr>
<th></th>
<th>WT-iBOAT</th>
<th>iBOAT</th>
<th>TRAOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial anomaly</td>
<td>35</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>Temporal anomaly</td>
<td>50</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2: Comparison of abnormal trajectory results.
7. Conclusion

Taking trajectory security as the research background, this paper provides an effective real-time risk monitoring scheme for detecting the anomalous trajectory of logistics vehicles. The main idea is to combine the spatial and temporal anomaly detection by using the time probability distribution of historical trajectory and taking the fixed position window as the detection fragment.

The results show that the algorithm can achieve higher accuracy, and the risk detection reaches 82.5% of the benchmark on our dataset, outperforming the traditional iBOAT algorithm (60%). At the same time, the false positive rate (FPR) of iBOAT is 0.031, and the FPR of our method is 0.044. Comparing the FPR values of the two methods, the higher FPR value of our method is caused by the supplemented temporal anomaly detection. However, we successfully found 22.5% more anomalies than iBOAT algorithm by using adjustment mechanism in temporal anomaly, so that the distribution of temporal anomaly results calculated by our method is more practicable as explained in Section 6.1. It is fairly acceptable that the FPR is increased by only 0.013 with 22.5% more anomalies correctly detected. Therefore, the result proves that our proposed algorithm can effectively detect the abnormal trajectories of urban freight vehicles in real time with not only higher accuracy, but also lower false alarm rate. This can provide more efficient, faster, and safer logistics express services for addressing unexpected trajectory anomaly events.

There are still some deficiencies in the current research. For example, when the historical trajectory data set is large, the real-time query of historical data will cost a lot of calculation, and the probability model will help solve this problem. Further research can establish probability model or neural network for different data to improve the speed of discovery from different start and end points.

In general, (1) the framework we use can be reasonably and widely applied to security service scenarios for urban freight vehicles. (2) For temporal anomalies, we propose a simple, fast, and efficient real-time temporal anomaly...
discovery method. (3) The proposed method provides idea for further research.

**Data Availability**

The trajectory data used to support the findings of this study may be released upon application to the Beijing Jiaotong University and the cooperative logistics company, who can be contacted at zhwei@bjtu.edu.cn.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Acknowledgments**

The research reported in this paper was supported by grants from National Natural Science Foundation of China (Grant no. 91746201) and National Key R&D Program of China (2018YFB1601600).

**References**


