Tracking and climbing behavior recognition of heavy-duty trucks on roadways

Lei Tang¹, Jingchi Jia¹, Zongtao Duan¹, Jingyu Ma¹, Xin Wang² and Weiwei Kong³

Abstract

The tracking and behavior recognition of heavy-duty trucks on roadways are keys for the development of automated heavy-duty trucks and an advanced driver assistance system. The spatiotemporal information of trucks from trajectory tracking and motions learnt from behavior analysis can be employed to predict possible driving risks and generate safe motion to avoid roadway accidents. This article presents a unified tracking and behavior recognition algorithm that can model the mobility of heavy-duty trucks on long inclined roadways. Random noise within the sampled elevation data is addressed by time-based segmentation to extract time-continuous samples at geographical locations. A Kalman filter is first used to distinguish error offsets from random noise and to estimate the distribution of truck elevations for different time intervals. A Markov chain Monte Carlo model is then applied to classify truck behaviors based on the change in elevation between two geographical locations. A heavy-duty truck mobility (HVMove) model is constructed based on the map information to apply the roadway geometry to the tracking and behavior recognition algorithm. We develop an extended Metropolis–Hastings algorithm to tune the parameters of the HVMove model. The proposed model is verified and evaluated through extensive experiments based on a real-world trajectory dataset covering sections of an expressway and national and provincial highways. From the experimental results, we conclude that the HVMove model provides sufficient accuracy and efficiency for automated heavy-duty trucks and advanced driver assistance system applications. In addition, HVMove can generate maps with the elevation information marked automatically.

Keywords

Heavy-duty truck, mobility modeling, spatiotemporal trajectory, elevation

Date received: 5 June 2019; accepted: 17 December 2019

Handling Editor: Rodolfo Meneguette

Introduction

Coasting and braking on long inclined roadways are one of the primary reasons for traffic accidents for heavy-duty trucks.¹–³ To improve truck climbing, we have to track and accurately predict the trucks’ movements while driving uphill and downhill. Ensuring the safety of drivers, trucks, and goods on roadways as well as the development of automated heavy-duty trucks can be facilitated by detailed analyses of truck motion and behavior. Moreover, such analyses also represent a primary foundation for evaluating environmental impacts,⁴,⁵ detecting air quality,⁶,⁷ reducing fuel consumption,⁸ and optimizing the transportation sector.⁹

¹School of Information Engineering, Chang’an University, Xi’an, China
²Department of Geomatics Engineering, University of Calgary, Calgary, Canada
³Shaanxi Skywin Internet of Trucks Information Technology Co., Ltd, Xi’an, China

Corresponding author:
Zongtao Duan, School of Information Engineering, Chang’an University, Xi’an 710064, Shaanxi, China.
Email: ztduana@chd.edu.cn
The specific motion of heavy-duty trucks depends on controllable tractive and braking forces as well as external forces arising from conditions such as road slope. Analysis of the dynamic features of trucks on continuous long upgrades is of particular importance because this facilitates the modeling of their mobility, thereby establishing the safety margin for surrounding moving trucks, particularly for trucks following in the rear. Tracking the trajectory of trucks and climbing behavior recognition are two important functions that facilitate the accurate modeling of truck mobility. Here, trajectory tracking is employed to estimate the dynamic features of trucks based on filtering approaches, while behavior recognition is a machine learning approach employed to recognize the specific motions of trucks based on the tracking results.

Many previous studies have focused on the tracking of vehicle trajectories on roadways. These tracking systems normally adopt algorithms that depend solely on the use of on-board sensors, such as video cameras, LIDAR, and radar for vehicle detection. However, these algorithms possess performance constraints based on the limitations of the sensors due to various factors, such as environmental impacts, the timeliness of multiple vehicle detection, and vibration. Tracking of heavy-duty trucks is further constrained when trucks travel through regions offering a narrow field of view and unexpected obstacles. However, these tracking constraints can be overcome by applying filters to constrain the detected motion of heavy-duty trucks. Therefore, it is possible to smooth the random noise introduced by motion-detection sensors and improve the accuracy of truck tracking and subsequent motion behavior recognition. In addition, trucks commonly travel along the same routes and make routine stops, such as for cargo loading and unloading, which leads to low-sampling-rate trajectories where the average time interval between consecutive sample points is greater than 10 s. As a result, the raw trajectories of trucks following highly structured routes in urban areas have fewer sample points than those of other vehicles such as taxies. Moreover, the quality of these truck trajectories cannot be substantially improved by the simple segmentation of each sample based on the spatial proximity between them.

Machine learning algorithms, such as the Bayes classifier, decision tree, and support vector machine, are used to infer the motion behavior of heavy-duty trucks on roadways. A truck motion model based on a deep neural network was developed by Wang et al. Although the deep learning approach enhanced the accuracy of the resulting truck mobility modeling, the model suffered from a lack of generality. For example, the complete driving operations of trucks could not be predicted when the trajectory data were incomplete. Thus, the application of deep learning to truck behavior recognition has been restricted by data sparsity. In addition, the precision of the deep learning approach depends on the number of kilometers traveled by a truck, which makes it difficult to recognize the long-distance mobility of trucks. An improved multi-mode hybrid automaton (MOHA) model was developed by Lin et al. for truck tracking purposes. The MOHA model extracts and clusters common state sequences from the temporal information of actual trajectory data based on a discrete event model and thereby identifies different truck behaviors. However, the MOHA model performs more poorly than the existing method due to high time complexity. In addition, a previous study introduced a unified framework for the tracking and behavior recognition of heavy-duty trucks under highway or constrained roadway driving conditions, but the algorithm did not consider the upgrade and downgrade motions of heavy-duty trucks. Furthermore, many previous studies have failed to consider the characteristics of heavy-duty trucks explicitly, such as long and fixed transportation routes and restricted traveling speeds and regions. These characteristics of heavy-duty trucks are distinct from, for example, the characteristics of taxies.

To address these issues, this article presents an algorithm for conducting the tracking and behavior recognition of moving heavy-duty trucks along long inclined roadways. A large-scale spatiotemporal trajectory dataset is used for measurements in the algorithm. The elevation information is extracted from the sampled trajectories. Based on the elevation information, a model, denoted as the HVMove model, is constructed, which adopts a logistic regression approach and produces the probability of truck motion based on the Markov chain Monte Carlo (MCMC) simulation. Tracking filters and behavior classification provide many benefits for conducting behavior recognition. Therefore, a Kalman-based filter serves as the basis of the HVMove model for generating the probability distribution of the elevation for different time intervals. The performance of the HVMove model is evaluated using real-world data.

The main contributions of this study are described as follows:

1. The proposed algorithm provides unified tracking and climbing behavior recognition of moving heavy-duty trucks;
2. The probability distribution of the climbing motions of heavy-duty trucks is modeled using a logistic distribution;
3. Maps with the elevation information are generated automatically from large-scale real-world trajectories.
Table 1. Summary of the GNSS data and geographic information of the areas analyzed in this study.

<table>
<thead>
<tr>
<th>Data</th>
<th>Datatype</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timestamp of received data</td>
<td>DateTime</td>
<td>20180402112300</td>
</tr>
<tr>
<td>GNSS longitude</td>
<td>long int</td>
<td>108950656</td>
</tr>
<tr>
<td>GNSS latitude</td>
<td>long int</td>
<td>34382758</td>
</tr>
<tr>
<td>GNSS elevation (m)</td>
<td>short int</td>
<td>346</td>
</tr>
<tr>
<td>GNSS speed (km/h)</td>
<td>short int</td>
<td>69</td>
</tr>
<tr>
<td>GNSS positioning mark</td>
<td>char</td>
<td>1</td>
</tr>
<tr>
<td>GNSS southern/northern latitude mark</td>
<td>char</td>
<td>0 or 1</td>
</tr>
<tr>
<td>GNSS eastern/western longitude mark</td>
<td>char</td>
<td>0 or 1</td>
</tr>
</tbody>
</table>

GNSS: global navigation satellite system. The total number of data records during the time period was 57,698.

The remainder of this article is organized as follows. Section “Data” presents the heavy-duty truck trajectory data used by the proposed method. Section “Labeling elevations in truck trajectories” presents the spatialtemporal trajectory model based on the Kalman filter, and section “MCMC-based truck motion model” describes the HVMove model based on two elevation features. Section “Performance analysis” presents the verification results, and section “Conclusion” concludes the article.

Data

The dataset employed in this work specifically includes sections of the Beijing–Kunming Expressway, G108 National Highway, and S311 Provincial Highway in the area of Shaanxi, and for Shanxi Province, China. General information regarding the time-stamped global navigation satellite system (GNSS) data within the 9-day period is summarized in Table 1. In our study, the data acquisition and positioning system are GPS and Beidou dual-mode navigation. The number of connected satellites is not less than four. The positioning error is 2.5 m that enables to fulfill the requirements for the high-sampling-rate trajectory. Each record of this dataset contains geographical location in the form of latitude and longitude, elevation, velocity, and time at each instance of heavy-duty truck activity, which includes traveling both up and down the slopes.

Truck coordinates are estimated from the records by on-board equipment based on a standard triangulation algorithm that provides an average coordinate error of 10 m. The spatiotemporal truck coordinate data represent the motions of heavy-duty trucks in connection with time and roadway. For example, a record from 10:25 a.m. to 11:47 a.m. on 1 April indicates a heavy-duty truck moving at low speed (<60 km/h) on a roadway with an inclined slope (elevation increasing from 344 to 542 m) for a long period of time (82 min). The data are intrinsically heterogeneous because the discrete approximate representations of geographical locations and elevations are derived using different sampling rates (e.g. every 500 m or 20 s). This work aims to provide a model that can support both the generation of origin-destination trips related to elevation and the identification of probability distributions for the motions of a heavy-duty truck for a particular period of a day based on the passive data employed.

Labeling elevations in truck trajectories

This section explains how the time-stamped GNSS records can be converted into individual trajectories with labeled elevations, which are then used to generate trip types for each heavy-duty truck. Due to the distinct sampling rates and poor data quality of the trajectories, we first perform the following data filtering: (1) data records with the same receiving times and geographic coordinates are removed; (2) a median filter is applied to adjust data records with similar coordinates, but with significantly different elevations; (3) lost elevation data points (56 records of all 57,698 records) are estimated using linear interpolation between consecutive spatiotemporal trajectory samplings; and (4) records that did not follow a strict time sequence due to factors such as variations in the geographical environment or signal quality were reordered into their proper sequences according to their sampling times.

Although the pretreated data records can describe the routes of heavy-duty trucks on an actual map, the random noise in the elevation data based on the GNSS trajectory segments should be reduced, which is discussed as follows.

Spatiotemporal trajectory segmentation and analysis

Trajectory segmentation is a significant step prior to engaging in further trajectory filtering. Here, segmentation can be conducted using the following three methods: (1) trajectory segmentation based on time intervals—here, if the time interval between two consecutive sampling points is greater than a given threshold, the trajectory is divided into two segments between the two points; (2) trajectory segmentation based on the spatial shape of the trajectory—in this method, the trajectory is segmented at the key points that maintain the trajectory shape; and (3) trajectory segmentation based on the semantics of a trajectory point—with this method, the trajectory is divided according to points such as a stationary or transitory point.

We combine the time-interval-based and semantics-based segmentation methods, where the semantics of the points are defined according to changes in the elevation of each truck (i.e. a truck traveling on an inclined roadway or on a declined roadway). Figure 1 presents a schematic diagram illustrating the conversion of daily
trajectory records to daily changes in elevation. Here, we first partition the trajectories according to every 24-h interval, detect locations where the truck is stationary (i.e., where the truck speed is 0 m/s), and then detect trips that occur between these stationary locations. Truck motions are generated over a specific time interval by first labeling all locations according to the relative changes in the elevation. These changes in elevation can be counted to determine the probabilities with which trips in each time interval involve increasing or decreasing elevation.

For example, a single truck in Figure 1(a) and (b) generates trips over the 2 single-day periods of observation, and these include 60 trips with an increase in elevation (indicated in blue), 76 trips with a decrease in elevation (marked in orange), while the remaining trips occur at a constant elevation (in white). These trips would be distributed across a single-day period based on the observation time of the stationary locations and the corresponding elevations, as shown in Figure 1(c) and (d). For example, the orange circle in Figure 1(c) indicates that 13% and 19% of all trips generated by heavy-duty trucks, respectively, exhibited an increase and decrease in elevations over the period of 6:00 a.m. to approximately 10:00 a.m. on a Saturday. Based on these visualizations from the prepared digital map, we can observe that the trips of heavy-duty trucks within the suburban and mountainous areas are more concentrated during the peak hours (6:00 to 8:00 a.m.) on Saturday and the late night hours (8:00 to 12:00 p.m.) on Monday due to their larger trip distances and less traffic. This procedure aims to generate a representative sample of trips to account for the travel choices of heavy-duty trucks within the suburban and mountainous areas of the region, as well as to label the elevation information on an actual map. Then, the elevation data noise was analyzed by detecting the short-term stationary points (SSPs) of trucks as follows.

**Definition 1 (SSP).** The SSP represents a segment of geographic data where a heavy-duty truck traveled at a speed of 0 m/s over a specified time interval. The extraction of an SSP depends on two scale parameters denoted as the time threshold ($\tau$) and the distance threshold ($\delta$). Formally, a single SSP can be obtained from a spatiotemporal trajectory characterized by points $(x, y)_i \rightarrow \cdots \rightarrow (x, y)_j$, that satisfy the conditions $\forall k \in [i, j], Dist((x, y)_k, (x, y)_{k+1}) < \delta, Int((x, y)_k, (x, y)_j) < \tau$, where $Dist(\cdot, \cdot)$ denotes the geospatial network-based distance between two points, $Int(\cdot, \cdot)$ is the time interval between two points, and $x$ and $y$ are the latitude and longitude, respectively.

Example SSPs extracted from pretreated trajectory data are shown in Figure 2. The distribution of random noise in the elevation data is more clearly observable
the data of a heavy-duty truck in the present work as $Tang et al. 5$

from the expanded data presented in the inset of Figure 2. Therefore, a Kalman filter is applied to smooth out short-term fluctuations in the time series data and thereby highlight the long-term trends of truck motions.

**Denoising by Kalman filter**

A Kalman filter is particularly advantageous for processing continuous noisy data points compared with median or mean filtering. With a Kalman filter, we define the state model for predicting the elevation of trucks as

$$H(k) = \Psi H(k-1) + Y W(k)$$  \hspace{1cm} (1)

where for the $k$th state at a single point in time, $H(k)$ is the elevation matrix, $W(k)$ is the noise matrix, and $\Psi = [1]$ and $Y = [1]$ are the state transformation matrices. The observation model used to obtain the elevation data from GNSS trajectories is then given for the $k$th state as

$$Y(k) = NG(k) + D(k)$$  \hspace{1cm} (2)

where $Y(k)$ is the observed elevation matrix based on GNSS data, $N = [1]$ is the observation matrix, $G(k)$ is the true elevation matrix, and $D(k)$ is the observed noise matrix. State and observed values can be connected. Therefore, the elevation in the $k$th state is estimated using equations (1) and (2) based on the last elevation value. The covariance matrix of the $(k-1)$th elevation is then updated

$$P(k|k-1) = \Psi P(k-1|k-1) \Psi^T + Q(k-1)$$  \hspace{1cm} (3)

where the parameters are determined using the Kalman filtering, and $Q$ is the covariance matrix of $W$. The Kalman gain $Kg$ is constructed for allocating the weights of the predicted and observed values as follows

$$Kg(k) = P(k|k-1)N^T[NP(k|k-1)N^T + R(k-1)]^{-1}$$  \hspace{1cm} (4)

where $R(k)$ is the covariance matrix of $D(k)$. According to the predicted value in equation (1) and the observed value in equation (2), the elevation of the $k$th state is obtained based on the weight calculated by equations (3) and (4). The matrices $P$ and $Kg$ for the corresponding state are updated as follows

$$H(k|k) = H(k|k-1) + Kg(k)[Y(k) - H(k|k-1)]$$

$$P(k) = (I - Kg(k))P(k|k-1)$$  \hspace{1cm} (5)

where $I$ is a unit matrix. Thus, the elevation data of each sampled point are obtained using equations (3)–(5).

**Elevation-labeled trajectory generation**

We define the trajectory model labeled by the elevation data of a heavy-duty truck in the present work as follows.

**Definition 2 (elevation-labeled trajectory).** An E-Tra is a sequence of time-stamped points $p_0 \rightarrow p_1 \rightarrow \cdots \rightarrow p_k$, where $p_i = (lat_i, lon_i, t_i, s_i)$, $(i = 0, 1, \ldots, k)$, in which $lat$ and $lon$ are the latitude and longitude, respectively, $t$ is the timestamp, $h$ is the elevation, $s$ is the speed, and the following conditions hold $\forall 0 \leq i \leq k, p_i(t_i+1) \geq p_i t$.

The E-Tra of heavy-duty trucks on an actual map is illustrated in Figure 3, in which each trajectory point is marked by the obtained elevation. The transport routes of the trucks and road conditions (such as gradient and length of slope) can be extracted from the map using the E-Tra model.

**MCMC-based truck motion model**

The elevation data optimized by Kalman filtering form the basis for extracting the features and for modeling the probability distribution of heavy-duty trucks
moving along long inclined roadways. The MCMC simulation was used in conjunction with the E-Tra model to determine the distribution of the features.

**Feature extraction**

Both the relative elevation difference (rED) and the sum of continuous elevation differences (i.e. the elevation difference summation—EDS) are used to track heavy-duty truck motions along an inclined slope. Therefore, they can be used as representations of truck behaviors. To extract the rED and EDS, we first model changes in the elevation between two consecutive trajectory points and obtain the duration for which a truck travels with the same type of motion, such as where the truck is continuously ascending or descending over a given time interval. Algorithm 1 was developed for extracting these features from the trajectories. Here, if the \( i \)th and \( (i + 1) \)th rED values are both positive or both negative, the \( i \)th rED value is accumulated in the EDS.

**HVMove model**

We considered the climbing behavior of heavy-duty trucks as a movement from a level roadway to an inclined roadway. The distribution is consistent with the shape of a logistic distribution. The aim of our proposed method is to model the probability distribution of truck climbing behavior denoted by \( M \) under a given rED. To achieve this goal, we utilize logistic regression to express the probability of \( M \) \( \{m_i, m_j\} \), as follows

\[
p(M = m_i|EDS) = \frac{1}{1 + e^{\beta(EDS)}}
\]  

(6)

where \( m_i = 0 \) denotes the state of a truck traveling on a level roadway, and \( m_j = 1 \) denotes the truck state of moving along an inclined roadway. We can estimate the parameter \( \beta \) by extending the MCMC simulation and thereby improve the HVMove model to fit well with the actual observation data, as follows

\[
p(M = m_i|EDS) = \frac{1}{1 + e^{\beta(EDS) + \alpha}}
\]  

(7)

The parameter \( \beta \) is used to determine the gradient of the model, while tuning parameter \( \alpha \) can alter the relative position of the distribution. The effects of the values of \( \beta \) (for \( \alpha = 0 \)) and \( \alpha \) (with different values of \( \beta \)) on the distribution of the HVMove model are illustrated in Figure 4(a) and (b), respectively. Based on the curves in Figure 4(a), we can observe that the slope

![Figure 3. Trajectories marked by elevation on an actual map.](image-url)
of the distribution is less than 0 when \( \beta < 0 \), while the slope is greater than 0 for \( \beta > 0 \). Figure 4(b) indicates that the distribution is offset to the left for \( \alpha < 0 \) and to the right for \( \alpha > 0 \). We note that the model reduces to equation (6) when \( \alpha = 0 \).

**Parameter tuning**

To find an appropriate model for describing the joint distribution of parameters \( \alpha \) and \( \beta \), we study the parameters under a two-dimensional Gaussian distribution. Accordingly, a Gaussian distribution was established with a mean value \( \mu = 0 \) and a standard deviation \( \sigma = 0.05 \) in Figure 5. Therefore, we generated 100 points between 0 and 5 at random by employing a Gaussian distribution illustrated as a thermodynamic histogram. Random sampling was performed in the hot areas to find an approximate solution for the HVMove model.

We then vary the values of \( \mu \) and \( \sigma \) in the Gaussian distribution to estimate the optimal parameters \( \alpha \) and \( \beta \) based on the MCMC simulation. To estimate the probability distribution of truck motion for given parameters, we consider the HVMove model as a 0–1 Bernoulli-variable-based representation, where the Bernoulli variable value of 0 denotes traveling on a level roadway and 1 denotes traveling on an inclined roadway. This is expressed as follows

\[
p(M = m_i|EDS) = Ber\left(\frac{1}{1 + e^{\beta\text{EDS} + \alpha}}\right)
\]  

(8)

We next show how to tune the parameters in the HVMove model step by step. We first use the Metropolis–Hastings algorithm to produce sample states of truck motion and evaluate the associated transition probabilities between two states with a generated Markov chain. The Markov chain structure of the HVMove model using the MCMC simulation is illustrated in Figure 6. Algorithm 2 describes the parameter tuning process, where an a priori Gaussian distribution \( N(\mu, \sigma^2) \) is used to establish the Markov chain of the distribution \( p(M|EDS) \). In Algorithm 2, a new state \( y \) is obtained from the Gaussian distribution, and the acceptance rate, \( a \), of this state is then calculated. In addition, comparing \( a \) with a random variable \( u \) from a uniform distribution \( \text{U} \) generates a new state and associated Markov chain.

After identifying the transition probability, we then apply the MCMC simulation to estimate \( \alpha \) and \( \beta \) according to the last state in each iteration. If the parameters fit with the actual data distribution, the current state is accepted; otherwise, the current state is rejected. Therefore, the sample set of each parameter can be obtained. By maximizing the likelihood of all samples, the optimal parameters can be learned, further building the HVMove model.

**Behavior recognition**

The proposed behavior recognition algorithm is given in Algorithm 3. First, SSPs are extracted. The elevation data are optimized by Kalman filtering. Then, the EDS
is obtained using Algorithm 1. The Markov chain of the distribution is established using Algorithm 2. Parameters $a$ and $b$ are obtained using the MCMC to build the HVMove model, which conducts heavy-duty truck motion behavior recognition.

**Performance analysis**

**Experimental setup**

We used the GNSS trace dataset of heavy-duty trucks presented in section “Data.” A total number of 70 trajectory sequences were obtained. The sequences that included fewer than 100 trajectory points were removed, and 48 sequences remained for fitting. Of these, the longest and shortest sequences comprised 2526 and 100 points, respectively. We then performed operations such as data cleansing and normalization to obtain the required elevation trajectories. Figure 7 investigates the effect of data processing. It compares the original SSP data of heavy-duty trucks with the processed data. It can be seen that the short-term fluctuations in the data segment declined significantly after Kalman filtering.

We randomly split the above 48 sequences into training and test sequences according to the ratio 7:3, that is, 70% of the data were used for training and the remaining 30% were used for testing. Measurement errors were identified using two different models. In the
first model, we analyzed the joint distribution of the time and the elevation for each truck. To find an appropriate model to describe this two-dimensional joint distribution, we can assume that a linear relationship exists between the time sample $T(t_1, t_2, \ldots, t_n)$ and the elevation sample $H(h_1, h_2, \ldots, h_n)$, that is, $T$ and $H$ are complied with a joint distribution $P(t, h)$. Supervised learning was used to determine the joint time and elevation distribution model, which is given as follows

$$h_n(t, \theta) = \theta_0 + \theta_1 t$$  \hspace{1cm} (9)

In the second model, the joint distribution of time and elevation was modeled based on a polynomial, as follows

$$h_n(t, \theta) = \theta_0 + \sum_{j=1}^{N} \theta_j t^j$$  \hspace{1cm} (10)

The testing set was used to evaluate the tracking performance of the two models by comparing the estimation error with the measurement error. Figure 8(a) presents the residuals of the two models, demonstrating that these models are suitable for defining the joint distribution of the elevation and time. Figure 8(b) illustrates the differences between the observations and predictions produced by the two models. The differences are generated using the mean absolute error (MAE) function. Because a straight line is applied to fit the data points, the slopes of the two lines represent the elevation estimation errors, where the larger the slopes, the smaller the errors. Figure 9 illustrates the uncertainties of the two models for four segments of SSPs based on the following expression

$$uA = \sqrt{\frac{\sum_{i=1}^{n} (\sigma_i - \bar{\sigma})}{(n-1)n}}$$  \hspace{1cm} (11)

where $n$ is the segment length. We can observe that the uncertainty of the linear model is less than that of the polynomial model when considering each segment with a given elevation range. Thus, we analyzed the errors of all data points from the trajectories using the linear regression method.

After calibration of the trajectory by offsetting the error using linear regression, we can label the ground-truth data in the dataset. If the elevation data were greater than or equal to a threshold, it was treated as positive samples (i.e. an increasing elevation), while the remaining data were treated as negative samples. To optimize the thresholds for ground truth labels, we performed 10-fold cross validation and identified the optimal threshold as 6.06 by minimizing the MAE of each sequence in all 48 sequences. We obtained 7119 sequences without SSPs. We consider these as our data-set and label their ground truth, that is, $(EDS(h_i + z_i \cdot h_2), L_m)^{7119}_{m=1}$, where $EDS()$ defines the EDS for a set of continuous trajectory points if and only if $h_2 - h_1 + z_i > 0$ as $i$ increases. The label is denoted as $L_m \in \{0, 1\}$. The labeled sequences are then input into the HVMove model for recognizing the climbing behaviors of heavy-duty trucks.

**Parameter study**

Figure 10(a) presents the distributions of $\alpha$ (top) and $\beta$ (bottom) with 5000 samples during HVMove model training. Considering that a greater number of iterations would increase the estimation accuracy in MCMC, we selected 500 $\alpha$ and $\beta$ values in the posterior section for calculating their probability distributions.

The recognition uncertainty would increase as the parameters become more widely distributed, and the overlapping between the sampled trajectories of heavy-duty trucks traveling on level and on inclined roadways has been investigated. Therefore, $\alpha$ and $\beta$ were averaged to determine the posteriori distribution of the motion behavior of trucks. The autocorrelations of the $\alpha$ and $\beta$ samples are shown in Figure 10(b), from which it can be seen that the convergent coefficients of $\alpha$ and $\beta$ maintained a decreasing trend and finally converged as the number of iterations reached about 20. Values of $\alpha = -5.069$ and $\beta = -11.507$ were then determined by taking the mean of all the parameter values when values are in the steady state.

**Performance of truck motion identification**

The motion behavior recognition results for trucks moving on level and inclined roadways are shown in Figure 11. The behavior recognition algorithm based on the HVMove model classifies motion behavior as climbing (CL) and flat road (FR) motion. However, we note from the figure that the HVMove model classifies the truck behavior in some regions as both CL and FR.
This is because the elevation variance is very small and the target truck changes its speed continuously. Therefore, FR behavior probabilities are dominant compared to CL behavior. The behavior recognition algorithm mostly classifies the scenarios as CL for the elevation changing scenario.
Conclusion

This study developed the HVMove model using spatio-temporal characteristics and pattern learning extracted from large-scale GNSS trajectory data for effectively modeling and predicting the ramp-climbing behavior of heavy-duty commercial trucks. First, an elevation-labeled trajectory, called as E-Tra model was established based on sampled trajectory data on a real map. The model provides positioning, time, altitude, and instantaneous speed of the truck, followed by the temporal segmentation of the trajectory data and optimization using Kalman filtering. The characteristics of the processed data were extracted and represented by logistic regression, for instance, the state transition, and the characteristic distribution of the ED was established using the MCMC simulation, which thereby determined the traveling mode characteristics of heavy-duty trucks. The HVMove model was finally established after determining the model parameters using the Metropolis−Hastings algorithm. In addition, the influence of the volume of sampled trajectory data on the predicted probability of ramp-climbing behavior was analyzed. The HVMove model can be integrated with a commercial in-car sensor system to track and identify the truck’s movements in time and to further predict and adapt the behaviors to minimize security risk while driving.

The model proposed herein was created primarily using the characteristics of elevation and time, whereas the influences of other characteristics such as truck speed, positioning marks, and fuel consumption on truck ramp-climbing behavior were not considered. In addition, only a single sampling method was used. Therefore, it is advisable to model and predict additional mobility features of heavy-duty trucks using different sampling methods and the characteristics of multi-source data in the future.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the Key Science and Technological Innovation Team of Shanxi Province, China (Grant No. 2017KCT-29) and the Key Research and Development Plan Project of the Shaanxi Province, China (Grant Nos 2019ZDLGY17-08 and 2019ZDLGY03-0901).

ORCID iD

Lei Tang https://orcid.org/0000-0001-5465-610X
References


