A Data-Driven Traffic Shockwave Speed Detection Approach Based on Vehicle Trajectories Data

Kaitai Yang
Department of Civil and Coastal Engineering
University of Maryland, College Park, MD, 20742
Email: kaitai74@umd.edu

Hanyi Yang, Ph.D
Department of Civil and Environmental Engineering
University of Hawaii at Manoa, Honolulu, HI, 96822
Email: hanyi@hawaii.edu

Lili Du* (Corresponding author), Ph.D
Department of Civil and Coastal Engineering
University of Florida, Gainesville, FL, 32611
Email: lilidu@ufl.edu

ABSTRACT

Traffic shockwaves demonstrate the formation and spreading of traffic fluctuation on roads. Existing methods mainly detect the shockwaves and their propagation by estimating traffic density and flow, which presents weaknesses in applications when traffic data is only partially or locally collected. This paper proposed a four-step data-driven approach that integrates machine learning with the traffic features to detect shockwaves and estimate their propagation speeds only using partial vehicle trajectory data. Specifically, we first denoise the speed data derived from trajectory data by the Fast Fourier Transform (FFT) to mitigate the effect of spontaneous random speed fluctuation. Next, we identify trajectory curves' turning points where a vehicle runs into a shockwave and its speed presents a high standard deviation within a short interval. Furthermore, the Density-based Spatial Clustering of Applications with Noise algorithm (DBSCAN) combined with traffic flow features is adopted to split the turning points into different clusters, each corresponding to a shockwave with constant speed. Last, the one-norm distance regression method is used to estimate the propagation speed of detected shockwaves. The proposed framework was applied to the field data collected from the I-80 and US-101 freeway by the Next Generation Simulation (NGSIM) program. The results show that this four-step data-driven method could efficiently detect the shockwaves and their propagation speeds without estimating the traffic densities and flows nearby. It performs well for both homogenous and nonhomogeneous road segments with trajectory data collected from total or partial traffic flow.

Keywords: Shockwave, Connected Vehicle, Clustering, Smoothening, Machine learning

1 INTRODUCTION

Traffic fluctuations resulting from different events, such as road construction, slow driver, traffic incidents, etc., will lead to traffic delays, capacity reduction, and extra emissions (Hegyi et al. 2008; Kerner et al. 1996; May 1990). Accordingly, a shockwave is a boundary between two neighboring traffic states (Hegyi et al. 2008) resulting from these events. The formation and dissipation of traffic shockwaves reflect the spreading of traffic fluctuation on a road segment. Therefore, detecting traffic shockwaves can help develop proactive traffic management such as ramp metering strategies (Qiu et al. 2010), traffic signal control (Lertworawanich & Unhasut 2021), adaptive speed advisory, etc. In addition, the detection of shockwaves helps identify highway bottlenecks and improve traffic operation, management, and control. Motivated by this view, there are quite a few studies in literature working on shockwave detection.

Early studies applied the data collected by fixed-pointed detectors (e.g., loop detectors, video cameras) for detecting traffic states variation (Dailey 1999; Hellinga 2002; Wang & Nihan 2003; Li, 2009) and then conducting shockwave analysis (Wang et al. 2016; Yang et al. 2014; Zanin et al. 2003; Liu et al. 2008). Briefly, under certain assumptions, traffic states such as density and speed on the target road section were estimated according to the data collected at fixed points on the highway. Then, the shockwave speed can be calculated for the target road section using the shockwave speed formulation (May 1990). This type of approach relies on the coverage of fixed-point sensors. It meets difficulty as the detection targets traffic on a long stretch of roadway or a network. Although high-dense point sensors have been implemented in some road segments, it is impractical to do it for the entire traffic network due to the high installation and management costs. On the other hand, the recent development of connected/probe vehicles (e.g., vehicles equipped with GPS or advanced communication and sensing technologies) offered comprehensive trajectory data resources for shockwave detection (Lu & Skabardonis 2007; Izadpanah et al. 2009; Khajeh-Hosseini & Talebpour 2019). Unlike point detectors mainly provide the aggregated traffic state data at a fixed spot, those probe vehicles provide comprehensive trajectory information, including location, speed, acceleration, and headway of individual vehicles in every short time interval (such as 0.1 seconds for the data used in this study). Thus, those trajectory data potentially can help us better understand the traffic features such as shockwaves. However, due to the limited penetration of probe vehicles, their trajectory data cannot offer an accurate estimation of traffic flow and density. To address this issue, some researchers introduced exogenous assumptions to validate the estimation of flow and density from trajectory data (Mao & Mao 2013; Seo et al. 2015), while others explored different approaches for shockwave detection. For example, the study by (Izadpanah et al. 2009) adopted wavelet transformation to smoothen vehicles' time-speed trajectories, then detected the shockwaves and their speed by searching and clustering the local minimums of each speed trajectory. Given local minimum speed is used as the key feature, this study could not detect the shockwave resulting from acceleration. Accordingly, the study by (Khajeh-Hosseini & Talebpour 2019) employed an iterative two-phase piecewise regression to identify shockwaves using vehicle trajectory data. Their approach involves three steps, including identifying the intersection points of shockwaves and vehicle trajectories, shockwave differentiation, and shockwave determination. It was noticed that this approach is very sensitive to the outliers that are mistakenly found in the first step. The study of Elfar et al. (Elfar et al. 2018) found that shockwaves often occurred in places where vehicle speed standard deviation increases significantly while identifying the shockwave traves and speed is not the main focus of this study. Overall, shockwave detection is still an open research question, and vehicle trajectory data provided researchers with new opportunities for detecting shockwaves and estimating their propagation speeds. Inspired by this view, this paper develops a four-step data-driven approach that seeks to detect all the shockwaves in the traffic stream during a detection window in real time using vehicle trajectory data. The data-driven approach is recursively implemented at each time step with a uniform look-back detection time window so that we can capture the shockwave propagation in real-time. Below we present the challenges and the methodology contribution, which integrates traffic flow theory, statistics, and machine learning methods.

First, according to the study Elfar et al. (Elfar et al. 2018), our data-driven approach takes the local increase of speed standard deviation as an indicator of a vehicle experiencing a shockwave. Accordingly, we want to screen each vehicle's trajectory to locate those turning points where traffic speed has significant changes. However, the actual trajectory data often involve noises (e.g., speed variation due to spontaneous driving behavior to obtain a better local driving experience), which cause spikes in trajectory curves and thus affect turning point detection. To address this issue, this study considered each trajectory curve as a time-series signal and applied Fast Fourier Transform (FFT) to denoise the raw trajectory data, e.g., to filter out those spikes before we detect turning points, given its proven performance in smoothening discrete data. On the other hand, we noticed that the speed standard deviation is not an effective/sensitive indicator to catch the turning points when traffic is highly congested. This is because vehicles are forced to travel at a low speed, and its speed variation is relatively mild even though it experiences a shockwave. To address this difficulty, this study detects the occurrence of a shockwave when noticing an extremely low local traffic speed.

Once turning points are detected, theoretically, we can find the shockwave propagation speed by connecting those turning points resulting from the same shockwave (see **Figure 1** (a)). However, this is not trivial. We need to address several challenges when real-time collected field trajectory data are used. First, we noticed that multiple shockwaves may co-exist in a time-space area and each of them causes turning points on vehicle trajectories as shown in **Figure 1** (b). Then the difficulty is how to accurately recognize which turning points belong to which shockwaves. Considering turning points from the same shockwave demonstrate certain patterns such as linear shape, our data-driven approach first separates turning points into clusters by an unsupervised machine learning algorithm (e.g., DBSCAN). We further notice that the machine learning approach cannot differentiate the turning points belong two different shockwaves while they are very close in a time-space area. Thus, this study develops an extra step involving traffic flow features. More exactly, we label each turning point with a positive or negative sign when it results from a shockwave causing acceleration or deceleration. Last, assuming each shockwave holds constant speed (i.e., turning points form a linear curve) in a short time interval, one-norm linear regression is designed to minimize the influence of noises that are not removed by the prior steps and estimate the propagation speeds of detected shockwaves.

This four-step data-driven approach is implemented as a recursive algorithm so that we can detect shockwaves in real-time. Namely, at each time step, we look back the data collected within a time window and detect the shockwaves and then move forward one time step for each recurve to involve real-time trajectory data. The performance of the data-driven method is tested by using the data collected on the I-80 and US-101 Freeway from the Next Generation Simulation (NGSIM) program (Alexiadis et al. 2004). The experiment results showed that this method could efficiently detect turning points and handle noise, and all the shockwaves were successfully identified and estimated. Its performance in identifying and estimating shockwaves is not affected as the penetration rate of the connected vehicle penetration rate reduces from 100% to 12.5%. Moreover, this method showed robust performance even for non-homogeneous traffic conditions, which is usually a required assumption for many existing studies.

The remainder of this paper is organized as follows. The second section formally defines the problem. The third section presents the data-driven method in technical detail. The fourth section evaluates the performance of the proposed method with the I-80 and US-101 datasets collected from the NGSIM program. The last section concludes this study with future research work.

2 METHODOLOGY

This study considers a road section of the length of L, on which a proportion of the traffic flow are probe vehicles that can provide trajectory data including locations (x), speed (v), and acceleration (a) at discrete sample timestamp $t \in \mathbb{Z}_+$ with a uniform time interval (Δt). Then, each data point, P, can be represented by a tuple, $P = (t, x, v, a)$. Our recursive data-driven approach seeks to detect the shockwaves on the road section L at each time step t using the trajectory data collected in a look-back time window

T(equal to 1 minute in our experiments). The computation time is much smaller than the look-back time window. In this way, we make a data-driven approach that can detect and update shockwaves of the closed road segment in real-time according to newly collected trajectory data at each time step. **Figure 2** illustrates this real-time recursive shockwave detection at time step t_1, t_2, t_3 with a look-back time window T.

The following sections will develop the four-step data-driven method applied at each time. Each step of the data-driven approach addresses one of the research challenges. Briefly, we first smoothen the trajectory of each vehicle, then take the standard deviations of the smoothed speeds as inputs to identify turning points presented in vehicle trajectories. After that, we classify the detected turning points by the unsupervised machine learning method, DBSCAN, combined with the characteristics of shockwaves. Last, one-norm linear regression is adopted to estimate the propagation speeds of each shockwave. Below we introduce the technical detail for each step of the data-driven approach.

2.1 TRAJECTORIES SMOOTHENING

Step 1 of the data-driven method seeks to smoothen the trajectory data by FFT. The idea behind this instrument is the observation that vehicle trajectories usually involve many minor speed fluctuations resulting from individual drivers' spontaneous actions to obtain a better local driving experience. These traffic fluctuations are not indicators for shockwaves and thus are considered as Type 1 noises. They present as the spikes in the curve of the trajectories (**Figure 3** highlights them in red circles). Type 1 noise affects affect the values of speed standard derivation and then the accuracy of the shockwave detection and also introduces extra computation load. To address this issue, this study considers trajectory data as a signal and smoothens vehicle trajectories by Fast Fourier Transform (FFT) shown in **Equation (1)** and **Equation (2)** since FFT (Brigham, E. O., & Morrow, R. E. (1967) has been proven to be a good method for obtaining the major trend of a time series data. It is expected to denoise the speed data to mitigate the effect of spontaneous random speed fluctuation.

$$\hat{v}[k] = \sum_{n=0}^{N-1} e^{-i\frac{2\pi}{N}nk} v[n],\ n = 0, 1, \dots, N-1 \tag{1}$$

$$v[n] = \frac{1}{N}\sum_{k=0}^{L-1} e^{i\frac{2\pi}{N}nk} \hat{v}[k],\ n = 0, 1, \dots, N-1 \tag{2}$$

where k represents the frequency of the transformed data, $v(n)$ denotes the original data, $\hat{v}[k]$ is the data transformed by FFT, and N is the total number of points.

To implement Fast Fourier Transform (FFT) approach, a vehicle's speed data, along with the time stamps was treated as a piece of signal with time and magnitude components consisting of N points. Specifically, **Equation (1)** transfers each vehicle trajectory $v(n), n = 0, 1, \dots, N-1$ from time domain to spectral domain ($\hat{v}[k]$). We noticed that those spikes resulting from spontaneous driving behavior correspond to the points with high frequency. Thus, we may drop out those noises and capture the main trend of the speed curve by only keeping those points with frequencies smaller than the threshold ℓ (set as 15 in our experiments according to our experiences working on the historical data). After that, the major information is transformed back to the time domain by **Equation (2)** (Parsons et al. 2000) so that we can obtain smoothed speed trajectory data. The influence of noises is largely mitigated, as shown in .

It should be noted that FFT cannot completely remove Type 1 noise. This is because mild shockwaves lead to similar speed fluctuations to the vehicle's spontaneous actions. There is no theoretical value of the threshold ℓ to strictly differentiate major speed trends and minor speed fluctuations resulting from spontaneous driving behavior. By understanding this point, this study makes the other three steps of our method have the capability/capacity to deal with the data with the errors. We will demonstrate those technical details in the following sections. The **algorithm I** below present the procedure of Step 1.

Algorithm I:

Step 1: Sorting the speed information according to time and vehicle ID. Taking the time series data for each vehicle as a signal.

Step 2: For each vehicle, implement Fast Fourier Transform to the time-speed trajectory.

Step 3: Keep the result obtained in Step 2 with lower frequencies and remove others.

Step 4: Implementing the inverse Fast Fourier Transform to the results obtained in Step 3 to obtain the smoothed vehicle trajectories.

2.2 TURNING POINTS IDENTIFICATION

Recall a vehicle's trajectory will present a significant speed change at turning points, which indicates where the vehicle crosses a shockwave. Therefore, local speed standard deviation has been selected as an indicator to identify turning points and further detect shockwave traces. However, our study noticed that when traffic is highly congested and traffic small is below 5mph, the speed standard deviation cannot serve as a good indicator since vehicles travel at such low speed, and they won't have significant speed changes even though they experience a shockwave. Consequently, Step 2 of our data-driven method takes both local speed standard derivation and low speed as indicators to detect the turning points. We explain the technical details as follows.

First of all, when traffic speed is above 5mph, this study identifies the turning points on the vehicle trajectories by the local speed standard deviation. Specifically, for each vehicle trajectory, starting from the first data point, this method takes a data set involving consecutive m data points to calculate the standard deviation (σ_{ij}) by **Equation** (3). The process moves forward on the data point each time. Namely, we first examine the standard deviation within the data set [1, m], and then sequentially move forward to do the same calculation within the data set [2, m+1], [3, m+2] until [N-m+1, N]. When the standard deviation of a data set is larger than the threshold ($D1$), we take the last data point in this set as the potential turning point since it is the new data point involved in this set as compared to its immediate proceeding data set and causes significant impacts on the speed standard derivation. The method above is implemented by Algorithm II.

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{m}(x_{ij} - \mu_j)^2}{m}}\ , \ \ j = m, m+1, \dots , N \qquad (3)$$

where x_{ij} represents the ith point within the window j, and μ_j is the mean of the window j.

Algorithm II:

Step 1: For each vehicle, remove the trajectory if $m > N$

Step 2: For each trajectory, starting from the first point to calculate the speed standard deviation of each window, j.

Step 3: For each trajectory, if $\sigma_j \geq D1$, then the last point of the corresponding window is regarded as the turning point.

Clearly, the window length, m**,** plays a key role in the calculation of the standard deviations. A larger window length may decrease the detection sensitivity and then reduce the shockwave detection accuracy. However, if the window length is too short, it will introduce many noises even though the trajectory data has been smoothened by the first step. The threshold $\boldsymbol{D1}$ also plays a significant role here. A large threshold will miss important turning points while setting a small threshold could lead to a messy result with many noises. This study suggests determining the optimal window length and threshold by the historical trajectory data. What is more, a shockwave is the boundary of two different traffic states. A vehicle that experiences a shockwave (i.e. going from one traffic state to another) will theoretically generate a single turning point on its trajectory if the vehicle can do a speed change immediately (see Figure 1 (a)). However, traffic states in reality do not have a clear boundary; when a vehicle goes through a shockwave, its speed change is conducted over a time period. As a result, using the indicator of speed standard deviation will find multiple turning points for one experience of a vehicle going through a shockwave (Figure 1 (b)). Consequently, our study noticed that directly using all these detected turning points makes our algorithm in Step 3 very sensitive to some parameter settings. Therefore, the turning points detected from the same vehicle within consecutive time points are aggregated as one point in the temporal-spatial diagram to indicate the impact of a shockwave on a vehicle. Accordingly, the coordinates of the new points are the means of the time and distance.

When traffic speed is below 5mph[1], our studies showed that speed standard deviation is not a sensitive indicator to finding the turning point or detecting shockwaves' impact on vehicle trajectories since vehicles traveling at such a low speed tend to change speed mildly even though they experience shockwaves. On the other hand, such low traffic speed in a local area indicates the occurrence of traffic congestion and its propagation (i.e., shockwave propagation). Therefore, this study considers a vehicle trajectory point as a potential turning point if its local velocity $v_h(t)$ is smaller than the threshold ($D2$). A sequence of low-speed points in a time period on a vehicle's trajectory indicates this vehicle experiences a mild speed deceleration and then acceleration. We use the first point as the turning point of a shockwave leading to speed deceleration and the last point as the turning of another shockwave leading to speed acceleration.

2.3 TURNING POINT CLUSTERING

The turning points reflect the impacts of the shockwaves on vehicle trajectories in a time-space area. A shockwave may affect a sequence of vehicles and generate a sequence of turning points. Theoretically, connecting those turning points gives us a trace of the shockwave (see the illustration in Figure 1 (b)). However, this is not trivial using trajectory data collected in the field. Mainly, Step 2 only gives us many turning points but cannot tell how many shockwaves co-exist and which turning point belongs to which shockwave on the road segment in a detection time window. On the other hand, our domain knowledge indicates that each shockwave will have its trace with different slopes or occur in different temporal-spatial zones. Accordingly, those turning points belonging to one shockwave will often form a cluster in the temporal-spatial detection area. Consequently, Step 3 of our data-driven method uses DBSCAN to cluster turning points found in Step 2, considering each cluster of turning points demonstrates a shockwave trace with constant speed. DBSCAN is selected for two reasons as follows.

First, multiple shockwaves with different speeds likely co-exist in the detected temporal-spatial diagram. However, we do not know the number of shockwaves. Therefore, a clustering method such as DBSCAN that can adaptively determine the number of clusters is preferred. Second, the turning points generated by Step 2 involve errors. This is because Step 1 cannot completely remove Type I noises, and Step 2 uses a probabilistic correlation between turning points and vehicle speed standard derivation to identify turning points. As a result, we will end with a turning point set with fake points (Type 2 noises), which do not indicate shockwave impacts on vehicle trajectories. Thus, the clustering algorithm that is adaptive to noises, such as DBSCAN, is preferred. Moreover, OPTICS is similar to DBSCAN but with a higher computation load. Moreover, OPTICS only produces a reachability distance plot and needs this study to cluster those turning points accordingly manually. Thus, it is not preferred. To be noted, this study does not intend to claim that DBSCAN is the only/best clustering solution for this project, but it does fit our

[1] This threshold is set offline based the traffic data in our experiments. Our approach is not very sensitive to this threshold. It can some values around 10 mph based on how you define the slow traffic in the applications.

requirements and performs well. Below we introduce the technical details for implementing DBSCN in this study. The DBSCAN algorithm was first introduced in (Ester et al. 1996), and the key idea of this method is to find the areas whose densities reach a threshold and thus are separated from other areas with lower densities without predefining the number of clusters. Specifically, there are two key parameters that need to be specified in this algorithm: Eps and $MinPts$. Here, Eps represents the distance measurement for checking the density in the neighborhood. The Euclidean distance is employed in this paper. In addition, $Minpts$ denotes the minimum number of points a neighborhood must contain. The DBSCAN can classify these turning points and automatically remove/drop the noise points based on the selected parameters for $Minpts$ and Eps. Specifically, DBSCAN method divided points into three different types, Core point, Border point, and Noise point. The points which have more than $Minpts$ neighbors within the radius Eps are considered as Core points. Those points that contain less than the $Minpts$ data points within the radius Eps but have at least one core point will be classified as Border points. All the rest points are treated as Noise points and will not be used for clustering. Therefore, the clustering approach is able to drop the outliers in the data set. Taking **Figure 4** as an example, the radius Eps is represented by circles. Point A, B, C are Core points as they meet the requirements of $Minpts$. Point D is a Border point and point E is a Noise point.

To implement this clustering algorithm for this study, scaling the coordinates and choosing proper values of Eps and $MinPts$ are critical. Large values of Eps and $MinPts$ cannot differentiate shockwaves, while small values of Eps and $MinPts$ may separate turning points belonging to one significant shockwave into multiple clusters and keep minor shockwaves and bring difficulty for further analysis. Existing studies have investigated many solutions to address this dilemma. This study scaled the time axis by 20 times and adopted the suggestion of (21) and made $MinPts$ equal to 4 since it has been shown a proper selection for 2-dimensional data. For the value of Eps, we follow the instruction of (21), and found that DBSCAN performs similarly under different traffic scenarios (see and) when we set the $MinPts$ equal to 400-700 but affect the computation load if it is larger than 400. Therefore, our experiments set $MinPts$ equal to 400, but only raise it to 600 when we test the performance of our method under very low data penetration, such as 12.5%. This uniform parameter setting also facilitates the application of our approach to conducting real-time shockwave detection. On the other hand, using the aggregated turning points in step 2 also promotes the generalizability of using DBSCAN with such uniform parameter settings in real-time applications.

Moreover, our experiments showed that only using DBSCAN cannot solve the problem perfectly since DBSCAN will cluster the turning points of two different shockwaves into one group when they are close to each in the temporal-spatial area. This is because DBSCAN clusters point according to the local density of points. When two shockwaves happen closely with a similar speed in a time-space area, their corresponding turning points are close. This often happens in traffic flow. For example, when an incident occurs on the road segment, traffic is blocked and will likely generate a backward shockwave, representing the formation of congestion propagating upstream. Once the incident is cleared quickly, the blocked traffic will restore to normal (i.e., vehicles accelerate to normal speed), and it will lead to another backward shockwave, which clears the traffic congestion. Consequently, these two shockwaves can be very close in a time-space area. Accordingly, their turning points stay close. DBSCAN does not work efficiently under this scenario. To address this problem, we noticed that the first backward shockwave makes vehicles decelerate, but the second backward shockwave makes vehicles accelerate. Even though these two shockwaves might be close in the time-space area, there is a chronological order between them. Therefore, the turning points representing one shockwave should not be included in the other. Therefore, we can first separate those turning points belonging to two different shockwaves by the sign of the acceleration of vehicle trajectories in the data before we run DBSCAN.

2.4 ESTIMATING THE PROPAGATION SPEED OF DETECTED SHOCKWAVES

Step 4 of this data-driven method seeks to find the shockwave speed. Specifically, we consider each cluster of turning points represents a shockwave with a constant speed. Then, each shockwave in one turning point cluster could be approximately treated as a linear shape on the target temporal-spatial area. Accordingly, using simple linear regression we can find the best-fitted shockwave curve with the format shown in **Equation (4)**. However, one limitation of simple least-square linear regression is that it is sensitive to outliers. This is because the parameters of the regression line are estimated by minimizing the 2-norm distance, i.e., the squared distance (see **Equation**).

$$\mathcal{L} = c_g + f_g * t, \tag{4}$$

where the dependent variable is distance, $\mathcal{L}$, and the independent variable is time, t. f_g represents the propagation speed of the target shockwave and c_g is the intercept which has no significant meaning in this project.

$$\sum_{q=1}^{w_g} (\tilde{\mathcal{L}}_q - \mathcal{L}_q)^2 \tag{5}$$

where w_g represents the number of total observations of the gth shockwave, $\tilde{\mathcal{L}}_q$ is the predicted value of the qth point in the target shockwave, and the observed value is denoted by $\mathcal{L}_q$.

To be noted, the DBSCAN algorithm could not perfectly separate turning points even though Step 3 introduces several instruments to promote the accuracy of the solution. In the meantime, Step 1 and Step 2 will cannot completely remove Type 1 and Type 2 noises due to the limits of the data-driven approaches. Consequently, the turning points in each cluster involve errors (named as Type 3 noises) which are either fake turning points generated by Step 2 or incorrectly clustered by Step 3. Those points cannot be statistically identified as outliers but could skew the estimation significantly. Mainly, they are not resulting from data collection mistakes and have the same physical meanings as other points. They are identified as noises merely because those nodes are close but not close enough to the major turning points of the corresponding shockwaves. Statistical outlier identification approaches cannot remove it either. As a result, the DBSCAN algorithm cannot filter them out, but they will affect the shockwave speed estimation. Also, the turning data in a cluster does not satisfy the assumptions of the 2-norm linear regression. For example, we noticed that the errors do not follow a normal distribution with constant variance. By recognizing these issues, using a method that is robust to this type of outliers/noises and does not require these residual assumptions will benefit this study more. Among various candidate approaches, the least absolute deviation (LAD) regression (one-norm regression) is employed to get a robust estimate since it has been proved that LAD regression performs better for resisting the influence caused by the outliers (Cade & Richards 1996) than general linear regression. **Equation (6)** shows that the one-norm regression minimizes the absolute value of the difference between the predicted value and the observed value, i.e., the one-norm distance.

$$\sum_{q=1}^{w_g} \left|\tilde{\mathcal{L}}_q - \mathcal{L}_q\right| \tag{6}$$

For the new model, the response and dependent variables are still the same. **Figure 5** shows one case in our study. It demonstrates that one-norm regression can capture the trend of the majority of turning points in a cluster and is much more robust to outliners than 2-norm linear regression.

3 NUMERICAL EXPERIMENT

This study evaluates the performance of the four-step data-driven method for detecting shockwaves by the numerical experiments built upon two trajectory datasets collected from the NGSIM project: The I-80 dataset and the US-101 dataset. The data was transformed from the video recordings, and the information of each vehicle was reported every 0.1 seconds. The parameter setting for all the cases is described as follows. To detect the potential turning points on the trajectories, we set $m = 30$ (i.e., 3 seconds), and the speed standard derivation threshold (D1) for recognizing a potential turning point was set as 2.5, after testing different options. The Eps is set to be 400 and the $Minpts$ is set to be 4 (the time axis is scaled by 20 times).

The experiments are conducted on a computer with Intel Core i9-9900K and Nvidia RTX 2070 super. For a 15-min look-back detection horizon, the data smoothing of Step 1 takes around 1s and the computing time for identifying turning points in Step 3 is less than 0.1s. The DBSCAN of Step 3 takes less than 0.1s and there is almost no waiting time for estimating the propagation speed using LAD regression. Therefore, the computation performance satisfies the real-time application. Below introduces the testbed, experiments, and main results in detail.

3.1 PERFORMANCE TEST BASED ON THE DATASET COLLECTED ON INTERSTATE FREEWAY I-80

The detailed vehicle trajectory data was collected on the eastbound I-80 in Emeryville, CA on April 13, 2005. The study segment contained five freeway lanes and a high-occupancy vehicle (HOV) lane and was approximately 1,640 ft long. Since the traffic state on the HOV lane is relatively stable, this study adopted the reconstructed trajectory data collected on lane 2 of I-80 (Montanino and Punzo 2013), from 4:00 p.m. to 4:15 p.m. to demonstrate the trajectories (**Figure 6** (a)). The slope changes of those trajectories indicate the traffic state changes by crossing shockwaves. Through visual observation, we can see that there are two major shockwaves, and each of them contains a backward shockwave forming the congestion and a backward shockwave clearing the congestion.

The turning points detected by Step 2 using the speed standard deviation as an indicator and the corresponding turning point clustering result of Step 3 are shown in **Figure 6** (b)-(d). Recall we split the turning points into two groups representing the shockwave causing traffic acceleration or deceleration separately. Correspondingly, the colored area in **Figure 6** (c)(d) represents their detected tuning point clusters, in which **Figure 6** (c) indicates the deceleration turning points clusters and **Figure 6** (d) illustrates the acceleration turning points clusters. To be noted, the noise points (black color) were removed for shockwave speed estimation in the next step. Similarly, the turning point result of Step 2 using the magnitude of speed as the indicator and the corresponding clustering results from Step 3 are shown by **Figure 6** (e)-(g). **Figure 6** (h) combines all the turning point clusters from **Figure 6** (c), (d), (f), and (g) and explores the shockwave traces estimated by the LAD regression. Specifically, the shockwave traces estimated from the turning points in **Figure 6** (c)(d) is shown by the orange and green lines. Given those turning points represent significant increases in speed standard deviation, the detected shockwaves lead to significant speed fluctuations. On the other hand, the shockwave traces detected from the turning points in **Figure 6** (f) and (g) are shown by the red and blue lines correspondingly. Given those turning points represent a sudden low or stop traffic in a local area, the detected shockwave traces indicate the stop-and-go scenarios. The two major shockwaves that could be detected by visual observation are shown by trace cluster 2 and trace cluster 5. Each trace cluster contains all 4 types of colored lines, which indicates that the vehicles' speeds suddenly decrease to stop and accelerate drastically when the congestion dissipates. It could be noticed that our data-driven approach detected more shockwaves than the visual observation shown in **Figure 6** (a). The numerical results are shown in **Table I** column 3, which will be discussed later.

3.2 PERFORMANCE TEST BASED ON THE DATASET COLLECTED ON US-101

The performance of the proposed data-driven approach is further evaluated on a more complicated scenario using the dataset collected on US-101 Freeway. Vehicle trajectory data was collected on southbound US 101 in Los Angeles, CA, on June 15th, 2005. The study segment was 2,100 ft long and contained five mainlines and an auxiliary line. This study takes the data collected on lane 1 from 7:50 a.m. to 8:35 a.m. as the input for shockwave detection.

The detected shockwave traces for the three 15-min intervals are shown in **Figure 7**. Note that this tested area is not a closed and homogeneous road segment since it includes both a mainline road segment and an auxiliary lane. The shockwaves present a periodical pattern and start from a similar location on the road. This phenomenon is counterintuitive as the data are collected from a freeway, and there is no traffic signal in the target area. One possible reason is that there is metering installed on the on-ramp adjacent to the testbed. Thus, the incoming flow from the on-ramp is periodic, which leads to these periodic traffic shockwaves on the freeway.

3.3 PERFORMANCE UNDER PARTIAL TRAJECTORY DATA

The experiments in Sections 3.1 and 3.2 evaluated the performance of the data-driven method under 100% trajectory data. It requires collecting every vehicle's trajectory in traffic flow. This is unrealistic in practice. Therefore, this study further tested the performance of the data-driven approach when trajectory data could only be collected from the partial traffic flow. To do that, we conducted experiments considering the penetration rate of the probe vehicles (e.g., CVs) as 50%, 25%, or 12.5%, and they were uniformly distributed in the traffic flow.

The results for the I-80 dataset are shown in **Figure 8**. Using the detected shockwaves and their slopes under a 100% penetration rate as the benchmark, we further evaluated the shockwave detection rate and the shockwave speed estimation errors under other scenarios with lower data collection penetration rates. The results are shown in **Table I**. By comparing **Figure 6** (h) and **Figure 8,** we observed that the data-driven approach could still detect most of the shockwaves (18/28) under the 50% penetration rate scenario; some shockwaves (such as trace cluster 3 and 4) hard to be visually recognized could still be identified (see **Figure 8** (a)). When the penetration rate dropped to 25%, some mild shockwaves only propagating in a small time-space area (such as trace cluster 6 in (a)) could not be detected (i.e., missing in **Figure 8** (b)) because there were not enough turning points to demonstrate the traffic pattern. When the penetration rate dropped to 12.5%, only the most obvious two clusters of shockwaves (trace clusters 2 and 5 in **Figure 6** (h)) were still successfully detected. Moreover, **Table I** shows that the majority (18/28) of the detected shockwaves under those low penetration rate scenarios presented a slope similar (error less than 10%) to what we found under 100% penetration; and only a few of the cases differed a lot compared to the 100% penetration rate scenario, such as trace cluster 1 and 3. This was because these shockwaves only propagated shortly. When the penetration rate dropped, the turning points used to estimate the slope of detected shockwaves trace were very limited, and further, the calculated slope was impacted significantly.

The results for using the US-101 dataset are given in **Figure 9** and the error percentages are shown in **Table II** columns 4-6. The number of detected shockwaves and the error percentages under different penetration rates showed a similar trend to what we observed in the I-80 case. Most of the shockwaves could still be detected for the 50% (34/43) and the 25% (25/43) penetration scenarios. And the total number of detected trace clusters was still consistent with visual observation for all 3 partial penetration rate scenarios except trace cluster 3 under the 12.5% penetration scenario. It was noticed that only some of the shockwaves identified from speed variations (orange and green lines) could be detected as the penetration rate dropped (such as trace cluster 6 in **Figure 7** (a)). On the other hand, for the shockwaves identified from the low magnitude of speed (red and blue lines), the difference between the 100% penetration rate and the partial penetration rate is relatively small. Overall, we claim that this data-driven approach is robust to CV penetration rates under most of the scenarios. It can detect major shockwaves with satisfying quality by only collecting a small portion of traffic trajectory data.

Overall, the experiment results illustrate that this data-driven approach is robust to data penetration rate. We explain the reasons from the merits of our methods as follows. With a lower probe vehicle penetration rate, there will be fewer turning points detected from the trajectory data. However, the DBSCAN can still recognize the turning point clusters corresponding to each shockwave, since those turning points can still present a similar and apparent pattern on the time-space diagram as the full penetration scenario. In other words, the approach used in this paper (i.e., DBSCAN) is mainly concerned with the distribution pattern of the turning point but not very sensitive to the number of turning points. On the other hand, this study used the LAD regression method to identify the trace/slope of the shockwave in each turning point cluster obtained from the DBSCAN. It is robust to the number of data also. More exactly, it will function well unless the turning points in each cluster can no longer demonstrate a clear trend. The number of turning points won't affect the results of the LAD regression significantly.

3.4 RESULT UNDER DIFFERENT SAMPLING FREQUENCIES

Except the partial penetration rate, the data sampling frequency may also impact the performance of the proposed approach. Thus, this study conducted the corresponding experiments to test it by varying data collection frequency from 0.1s, 0.5s, to 1s. **Figure 10**. demonstrates the results under 0.5s sampling frequency. We can see that the total number of shockwaves and the estimated shockwave speeds does not differ a lot compared to the 0.1s sampling frequency shown in **Figure 7**. Thus, we claim that the data-driven approach still works with a 0.5s sampling frequency. But our experiments noticed that the data-driven approach cannot function well if we further decreased the data collection frequency to 1s. This was because such sparse trajectory data cannot recognize sufficient turning points for shockwave trace detection. Up to now, 0.1s sampling frequency is popular in most GPS data collection equipment. Therefore, this approach should work well under normal GPS data collection.

4 CONCLUSION

This paper proposed a four-step data-driven algorithm implemented as a recursive algorithm for detecting shockwaves and their propagation speed on a closed road segment based on real-time collected vehicle trajectory data. The four steps include trajectory smoothening, turning points detection, shockwave clustering, and shockwave speed estimation. Specifically, Fast Fourier Transform was first adopted to denoise each trajectory curve for removing the spontaneous traffic variation. Then, we screen the speed data by measuring the standard deviation of the speed within a detection time window. A time window with a standard deviation larger than the defined threshold is treated as a turning point of the trajectory curve. Next, the DBSCAN method combined with traffic flow features is applied to properly cluster the turning points so that each cluster corresponds to a shockwave. Last, one-norm linear regression is used to identify shockwave curves and estimate their speed due to its good performance in mitigating the effect of the outliers. At each step, we integrate traffic flow characteristics with the main method to make our data-driven approach more applicable in reality. The performance of the four-step data-driven approach is tested by using the field trajectory data collected on the I-80 and US-101 Freeway. The results confirmed the efficiency and accuracy of this approach to detect shockwaves for both homogenous and nonhomogeneous road segments with trajectory data collected from either full or partial traffic flow. Therefore, we can apply this data-driven approach to detect shockwaves on the roads without sufficient fixed-point sensors but having a small proportion of trajectory data. In the meantime, we noticed that this method is not sensitive to shockwave propagation under slow traffic conditions since the turning points are detected based on speed variation. Our future work will make this approach more robust and efficient under various traffic conditions.

ACKNOWLEDGMENTS

This research is partially supported by the National Science Foundation awards CMMI-1901994, CMMI-2213459 and CNS-2124858. The authors would like to extend their gratitude to the reviewers and editor for their insightful comments, which have increased the quality of this paper.

AUTHOR CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: Dr. L. Du initiated this idea and supervised the whole study. Students K. Yang and Dr. H. Yang conducted the approach development, implementation, and data collection. All three authors drafted, edited, and reviewed the manuscript. They all reviewed the results and approved the final version of the manuscript.

REFERENCES

Alexiadis, V., Colyar, J., Halkias, J., Hranac, R., & McHale, G. (2004). The next generation simulation program. Institute of Transportation Engineers. ITE Journal, 74(8), 22.

Brigham, E. O., & Morrow, R. E. (1967). The fast Fourier transform. IEEE spectrum, 4(12), 63-70.

Cade, B. S., & Richards, J. D. (1996). Permutation tests for least absolute deviation regression. Biometrics, 886-902.

Dailey, D. J. (1999). A statistical algorithm for estimating speed from single loop volume and occupancy measurements. Transportation Research Part B: Methodological, 33(5), 313-322.

Elfar, A., Xavier, C., Talebpour, A., & Mahmassani, H. S. (2018). Traffic shockwave detection in a connected environment using the speed distribution of individual vehicles. Transportation Research Record, 2672(20), 203-214.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In Kdd (Vol. 96, No. 34, pp. 226-231).

Hegyi, A., Hoogendoorn, S. P., Schreuder, M., Stoelhorst, H., & Viti, F. (2008, October). SPECIALIST: A dynamic speed limit control algorithm based on shockwave theory. In 2008 11th international ieee conference on intelligent transportation systems (pp. 827-832).

Hellinga, B. R. (2002). Improving freeway speed estimates from single-loop detectors. Journal of transportation engineering, 128(1), 58-67.

Izadpanah, P., Hellinga, B., & Fu, L. (2009, January). Automatic traffic shockwave identification using vehicles' trajectories. In Proceedings of the 88th Annual Meeting of the Transportation Research Board (CD-ROM).

Kerner, B. S., & Rehborn, H. (1996). Experimental features and characteristics of traffic jams. Physical review E, 53(2), R1297.

Khajeh-Hosseini, M., & Talebpour, A. (2019, October). Back to the Future: Predicting Traffic Shockwave Formation and Propagation Using a Convolutional Encoder-Decoder Network. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 1367-1372). IEEE.

Lertworawanich, P., & Unhasut, P. (2021). A CO emission-based adaptive signal control for isolated intersections. Journal of the Air & Waste Management Association, 71(5), 564-585.

Li, B. (2009). On the recursive estimation of vehicular speed using data from a single inductance loop detector: a Bayesian approach. Transportation Research Part B: Methodological, 43(4), 391-402.

Liu, Z., Chen, Y., & Li, Z. (2008, June). Vehicle queue detection based on morphological edge. In 2008 7th World Congress on Intelligent Control and Automation (pp. 2732-2736). IEEE.

Lu, X. Y., & Skabardonis, A. (2007, January). Freeway traffic shockwave analysis: exploring the NGSIM trajectory data. In 86th Annual Meeting of the Transportation Research Board, Washington, DC.

Mao, R., & Mao, G. (2013, April). Road traffic density estimation in vehicular networks. In 2013 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 4653-4658). IEEE.

May, A. D. (1990). Traffic flow fundamentals. Prentice-Hall, Englewood Cliffs, NJ

Montanino, M., and V. Punzo (2013). Reconstructed NGSIM I80-1. COST ACTION TU0903 - MULTITUDE http://www.multitude-project.eu/exchange/101.html.

Parsons, S., Boonman, A. M., & Obrist, M. K. (2000). Advantages and disadvantages of techniques for transforming and analyzing chiropteran echolocation calls. Journal of Mammalogy, 81(4), 927-938.

Qiu, T. Z., Lu, X. Y., Chow, A. H., & Shladover, S. E. (2010). Estimation of freeway traffic density with loop detector and connected vehicle data. Transportation Research Record, 2178(1), 21-29.

Seo, T., Kusakabe, T., & Asakura, Y. (2015). Estimation of flow and density using connected vehicles with spacing measurement equipment. Transportation Research Part C: Emerging Technologies, 53, 134-150.

Wang, Y., & Nihan, N. L. (2003). Can single-loop detectors do the work of dual-loop detectors? Journal of Transportation Engineering, 129(2), 169-176.

Wang, J., Xie, W., Liu, B., & Ragland, D. R. (2016). Identification of freeway secondary accidents with traffic shock wave detected by loop detectors. Safety science, 87, 195-201.

Yang, D., Chen, Y., Xin, L., & Zhang, Y. (2014). Real-time detecting and tracking of traffic shockwaves based on weighted consensus information fusion in distributed video network. IET Intelligent Transport Systems, 8(4), 377-387.

Zanin, M., Messelodi, S., & Modena, C. M. (2003, September). An efficient vehicle queue detection system based on image processing. In 12th International Conference on Image Analysis and Processing, 2003. Proceedings. (pp. 232-237). IEEE.