Research Article

Extracting Origin-Destination with Vehicle Trajectory Data and Applying to Coordinated Ramp Metering

Cheng Zhang,1 Jiawen Wang,2 Jintao Lai,1 Xiaoguang Yang,1 Yuelong Su,3 and Zhenning Dong3

1The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China
2Department of Traffic Engineering, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, China
3AutoNavi Software Co., Beijing 100102, China

Correspondence should be addressed to Jiawen Wang; wangjw@usst.edu.cn

Received 8 January 2019; Revised 18 April 2019; Accepted 12 June 2019; Published 8 July 2019

Academic Editor: Zhi-Chun Li

Copyright © 2019 Cheng 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.

Ramp metering is an effective measure to alleviate freeway congestion. Traditional methods were mostly based on fixed-sensor data, by which origin-destination (OD) patterns cannot be directly collected. Nowadays, trajectory data are available to track vehicle movements. OD patterns can be estimated with weaker assumptions and hence closer to reality. Ramp metering can be improved with this advantage. This paper extracts OD patterns with historical trajectory data. A validation test is proposed to guarantee the same representativeness of vehicle trajectories and then implement coordinated ramp metering based on the contribution of on-ramp traffic to downstream bottleneck sections. The contribution is determined by the OD patterns. Simulation experiments are conducted under real-life scenarios. Results show that ramp metering with trajectory data increases the throughput by another 4% compared with traditional fixed-sensor data. The advantage is more significant under heavier traffic demand, where traditional control can hardly relieve the situation; in contrast, our control manages to make congestion dissipate earlier and even prevent its forming in some sections. Penetration of trajectory data influences control effects. The minimum required penetration of 4.0% is determined by a t-test and the Pearson correlation coefficient. When penetration is less than the minimum, the correlation between the estimation and the truth significantly drops, OD estimation tends to be unreliable, and control performance becomes more sensitive. The proposed approach is effective in recurrent freeway congestion with steady OD patterns. It is ready for practice and the analysis supports the real-world application.

1. Introduction

Ramp metering has been proved to be an effective measure to relieve freeways congestion. It regulates the traffic volume at entry ramps with traffic signals, to keep the mainline traffic operating at the desired level and avoid capacity drop [1]. Proper entry control reduces the probability of breakdowns on the mainline while also considering the throughput of other traffic [2].

Researches on ramp metering can be dated back to 1965 when Wattleworth first proposed a linear programming model to generate a fixed-time control scheme to alleviate peak-hour freeway congestion with static origin-destination (OD) information [3]. With real-time measurement available, reactive ramp metering strategies have been proposed. The best-known strategies are the demand-capacity (DC) strategy, the occupancy (OCC) strategy, and the ALINEA. DC strategy allowed ramp volume to make up the difference between the estimated capacity of mainline and measured upstream volume. OCC strategy calculated the metering rate by comparing the measured occupancy and the target value [4]. ALINEA was a feedback regulator, aiming to maintain capacity flow downstream of the merge area [5]. All these three strategies are local ramp metering based on simple logic. Field results showed their effectiveness [6]. Some extension versions were also proposed such as PI-ALINEA [7] and ITC-ALINEA [8]. They belong to local ramp metering strategies.
Coordinated ramp metering considers coordination among ramps. Cooperative algorithms, such as HELPER [9] and HERO [10], determined critical ramps whose downstream mainline sections are in most serious congestion and then force its upstream on-ramps to reduce metering rates accordingly. Bottleneck algorithms, such as FLOW [11], ZONE [12], and SWARM [13], calculated metering rates based on volume reduction in the mainline bottleneck. The above two kinds of algorithms are rule-based ones. The logic is like “IF traffic condition, THEN control scheme”, which is easy to implement.

Another kind of coordinated ramp metering control is optimal control. Started in the 1960s, the optimizing models inputted historical data and outputted static or time-of-day control schemes [14]. In the 1990s, real-time measurements were considered. The approach used a rolling horizon to get a dynamic optimal control scheme [15]. Optimal control can generate the best solution in spite of complex modeling and computing. Among them, linear programming was relatively easy to solve, so it had been first developed and implemented [16]. Nonlinear relationships of traffic systems had also been discussed and applied to ramp metering [17].

A majority of these researches were based on macroscopic traffic models, including Cell Transmission Model [18], Payne–Whitham model [19], and METANET [20]. They adopted models to describe the relationships among variables and simulate the evolution of freeway systems. Artificial intelligence was also employed to ramp metering. These techniques included fuzzy logic [21], neural networks [22], iterative learning [8], and other heuristic methods [23].

For most coordinated ramp metering, origin-destination (OD) demand is a vital input. Traditionally, OD matrices were derived from surveys, which were time-consuming and usually not precise enough due to subjective response and small sample size. With the development of traffic detection technology, many methods were proposed to estimate OD with different data sources. Generally, they can be divided into two categories—the fixed-sensor-based methods and the trajectory-based ones.

Fixed sensors include loop detectors, radar detectors, and video sensors. They collect traffic information like volume, occupancy, and speed at certain locations. A lot of researches were conducted to estimate OD with fixed-sensor data. One of the most primitive approaches is proportional distribution method [24]. It regarded that the OD volume on freeways is in proportion to the volume at entrance ramps. Obviously, this assumption was too strong. A gravity-based model was then proposed [25]. It assumed that the trip distances meet Gamma distribution, which implied that too long or too short trips are of little possibility. Also, estimated travel time distributions were applied to OD matrix estimation [26]. Besides, commonly used models include generalized least square model [27], the maximum entropy model [28], and Bayesian theory [29]. Since fixed sensors cannot capture traffic’s origin and destination, these methods had to introduce assumptions to estimate OD.

Trajectory data record the movement of individual vehicles and thus provide a more reliable data source for OD estimation. The first kind of data is returned by detectors on vehicles, such as floating car data collected by GPS equipment [30] and vehicle trajectory data returned by mobile navigation applications [31]. Since not all vehicles are equipped with GPS devices or not all travelers use navigation apps, this kind of data is sample data. Penetration rate influences estimation accuracy [32]. The second kind of data is collected by widely distributed fixed sensors, like automatic license plate recognition data [33] and Bluetooth and WiFi data [34]. Different from traditional fixed sensors, they identify each vehicle or traveler via an identification (ID), which makes trajectory reconstruction possible.

Nowadays high-quality trajectory data have become available. Previous researches have utilized trajectory data to estimate traffic state parameters such as OD flows [31], queue length [35], delay [36], and travel time [37]. Results showed that estimation accuracy was improved with trajectory data. If traffic parameters estimated by trajectory data could be applied to ramp metering, the control performance is supposed to be improved. However, limited research or practice has implemented it since high-quality trajectory data were not available. Low frequency (e.g., > 30 seconds of record interval) and low penetration (e.g., <10%) are usual situations. Thus, most existing ramp metering controls were still based on fixed-sensor data—typically loop detector data. The OD table was estimated by link counts [19]. Instead of estimating OD tables, some researches introduced off-ramp exit percentages to predict freeway flows [38]. Vehicles from different origins have the same probability to leave the mainline when they come to an off-ramp [39]. This assumption is not necessarily true. Others assumed that the OD flows were perfectly known and the studies were conducted under ideal scenarios [18]. However, these researches have a long distance from practice.

The contribution of this paper includes (1) proposing and implementing coordinated ramp metering with real trajectory data; (2) comparing the control performance of the proposed approach and the traditional method under different traffic demand intensity; (3) analyzing the influence of penetration rate on control performance and determining a minimum value. The remaining part of this paper is organized as follows: Section 2 introduces the bottleneck algorithm implemented in this research. Section 3 explains the process of weights calibration with trajectory data. Section 4 presents simulation experiments settings. Section 5 compares the experiment results. Section 6 analyzes the sensitivity on penetration rate. Section 7 concludes this paper.

2. Bottleneck Algorithm

The bottleneck algorithm is essentially a rule-based algorithm for coordinated ramp metering control. Ramp metering rates are calculated, respectively, on the basis of system capacity constraints and local conditions. The more restrictive metering rate is selected and finally adjusted to constraints including maximum queue length and minimum metering rate.

Figure 1 shows relevant variables of bottleneck sections and on-ramps. $i$ is mainline section index, $j$ is on-ramp index,
and \( k \) is the time step. A mainline section \( i \) is defined as a bottleneck if its average occupancy \( \rho_i \) is over the threshold \( \rho_i^{bres} \) (see (1)); at the same time the total entering flow (including inbound flow \( q_i^{in} \) and on-ramp flow \( q_i^{on} \)) is more than the total exiting flow (including outbound flow \( q_i^{out} \) and off-ramp flow \( q_i^{off} \)) (see (2)).

\[
\rho_i (k) > \rho_i^{bres} 
\]

\[
q_i^{in} (k) + q_i^{on} (k) > q_i^{out} (k) + q_i^{off} (k) 
\]

A bottleneck metering rate \( r_j^B \) is calculated if there exists a bottleneck on the freeway. On the basis of the metering rate \( r_j \) in the last time step, entering traffic is reduced to alleviate congestion in bottleneck sections. The volume reduction is determined according to the law of traffic volume conservation. For bottleneck section \( i \), its volume reduction \( q_i^{reduce} \) is the difference between the total entering flow and the total exiting flow (see (3)). Then the reduction is distributed to its area of influence, i.e., its upstream on-ramps according to weights \( W_{ij} \) (see (4)). The weights calculation is discussed in Section 3.4 in detail. When there are more than one bottlenecks, their areas of influence overlap. On-ramp volume reduction \( q_{ij}^{reduce} \) can be calculated from each bottleneck. The maximum volume reduction is used to calculate the bottleneck metering rate (see (5)).

\[
q_i^{reduce} (k + 1) = [q_i^{in} (k) + q_i^{on} (k)] 
- [q_i^{out} (k) + q_i^{off} (k)] 
\]

\[
q_{ij}^{reduce} (k + 1) = q_{ij}^{reduce} (k + 1) \cdot W_{ij} 
\]

\[
r_j^B (k + 1) = r_j (k) - \max \left[ q_{ij}^{reduce} (k + 1) \right] 
\]

Then, a local metering rate \( r_j^L \) is calculated on the basis of the local condition of on-ramp \( j \) using demand-capacity strategy. To be specific, if the mainline occupancy downstream of the on-ramp \( \rho_j \) does not reach the threshold \( \rho_j^{bres} \), the metering rate is the difference between capacity \( q_j^{cap} \) and mainline traffic flow upstream of the on-ramp \( q_j^{in} \); otherwise, it equals the preset minimum value \( r_j^{min} \) (see (6)).

\[
\rho_j (k) > \rho_j^{bres} 
\]

\[
q_j^{cap} - q_j^{in} (k) - \rho_j^{bres} + \rho_j^{bres} \leq \rho_j 
\]

\[
r_j (k + 1) = \max \left[ r_j^L (k + 1), r_j^{queue} (k + 1), r_j^{min} \right] 
\]

The occupancy threshold \( \rho_j^{bres} \) is corresponding to the state when the mainline reaches capacity flow.

\[
r_j^L (k + 1) = \begin{cases} 
q_j^{cap} - q_j^{in} (k) & \rho_j (k) \leq \rho_j^{bres} \\
\frac{r_j^{min}}{r_j^{min}} & \rho_j (k) > \rho_j^{bres}
\end{cases} 
\]

If on-ramp \( j \) is the nearest on-ramp upstream of bottleneck section \( i \), then the inbound flow upstream of the bottleneck section \( i \) equals mainline traffic flow upstream of the on-ramp \( j \) (see (7)). Also, the on-ramp flow of the bottleneck section \( i \) is the minimum of the flow corresponding to the metering rate and the flow corresponding to the vehicles waiting in the on-ramp \( j \) plus the arriving vehicles (see (8)).

\[
q_j^{in} (k) = q_j^{in} (k) 
\]

\[
q_j^{on} (k) = \min \left[ r_j (k) , d_j (k) + \frac{n_j (k)}{T} \right] 
\]

where \( r_j \) is the executed metering rate, \( d_j \) is the vehicle arriving rate, \( n_j \) is the initial vehicle number on the ramp, and \( T \) is the duration of one time step.

The system metering rate \( r_j^S \) adopts the more restrictive of the bottleneck metering rate \( r_j^B \) and the local metering rate \( r_j^L \).

\[
r_j^S (k + 1) = \min \left[ r_j^L (k + 1), r_j^B (k + 1) \right] 
\]

Then, the metering rate is adjusted to prevent the vehicle queue of each on-ramp exceeding the storage capacity \( N_j \). The vehicle queue in the next time step is the initial vehicles on the ramp plus arriving vehicles and then minus released vehicles (see (10)). Thus deduce the minimum metering rate to satisfy the queue limits \( r_j^{queue} \) (see (11)).

\[
n_j (k) + [d_j (k + 1) - r_j (k + 1)] T \leq N_j 
\]

\[
r_j^{queue} (k + 1) = d_j (k + 1) + \frac{n_j (k) - N_j}{T} 
\]

Besides, the metering rate should be no less than the preset minimum value \( r_j^{min} \). Thus, the final metering rate is the maximum of \( r_j^S, r_j^{queue} \) and \( r_j^{min} \).

\[
r_j (k + 1) = \max \left[ r_j^S (k + 1), r_j^{queue} (k + 1), r_j^{min} \right] 
\]
Finally, the metering rate is converted to green time $g_j$.

$$g_j (k + 1) = \frac{r_j^\text{sat}(k + 1)}{r_j^\text{sat}} \times C_j$$  \hspace{1cm} (13)

where $r_j^\text{sat}$ is the saturation flow (usually 1800 pcu/h) and $C_j$ is the signal cycle. Figure 2 presents the flow chart of the bottleneck algorithm described above.

3. Weights Calibration

3.1. Data Description. Data used in this research include trajectory data, loop detector data, and a corresponding road network map. Trajectory data is collected from a typical week, from 23 October 2017 to 29 October 2017, including about 13 million records. The indicators include a timestamp, vehicle ID, speed, location in terms of longitude, latitude, and road ID matched to the road network map. Vehicle trajectory data are collected by mobile navigation applications. They are a kind of crowd-sourced data. With the development of mobile Internet technology, a great number of drivers use mobile navigation applications such as Google Map and Amap during their trips. To provide real-time navigation service, these mobile applications frequently collect and update vehicle location and movement, as shown in Figure 3. Each vehicle is assigned an ID, with which its trajectory can be accurately reproduced. Mobile phones with these navigation applications are essentially mobile sensors. The vehicle trajectory data can be regarded as an improved version of floating car data with higher collection frequency and higher penetration.

Loop detector data cover the mainline and ramps of North-South Elevated Road in Shanghai, China. The duration is from 1 October 2017 to 31 October 2017. There are about 0.9 million records of traffic statistics in 5 minutes, including volume, occupancy, and speed. Detectors can be located to map via loop ID. Besides, the map contains information about the road network, such as road name, length, width, and number of lanes of the link.

3.2. Origin-Destination Extraction. Origins and destinations of trips can be extracted from vehicle trajectory data. Generally, trips are distinguished by different purposes. Vehicle trajectory is divided into multiple trips based on dwell time and dwell location [33]. In this research, the problem is discussed in freeway networks where vehicles normally keep moving. Also, trajectory is collected every 5 seconds on average, forming a chain of continuous records as shown in Figure 3. These advantages simplify the process of OD extraction. When a vehicle first appears in the freeway network, the corresponding record is regarded as the origin of a trip. The following records all belong to this trip until one record appears at a nonadjacent road or the time gap between this record and the previous sequence is large enough. Then, this record signifies a new trip. Since normally there does not exist trip purposes on the freeway, the OD extraction algorithm only needs to examine temporal and spatial continuity for trajectory records in the same trip.

Figure 4 presents a flow chart to extract OD pairs from trajectory data. First, all trajectory data are ordered by vehicle ID and timestamp. The algorithm compares two contiguous records. If they have the same vehicle ID, their time gap is less than a threshold (e.g., 10 minutes), and the path is continuous, then these two records belong to the same trip and the temporary destination is updated to the latter record; otherwise, the former record is the destination of this trip and the latter record is the origin of a new trip. Figure 5 shows an example of ordered trajectory data. They are divided into several trips through Figure 4. In this way, the origin and the destination of each trip and their corresponding location and time are extracted.

3.3. Validation Test. In order to verify that the OD pairs extracted from trajectory data are representative samples, we count the number of vehicles passing through a link within a given time interval from trajectory data and count the flow volume of the same link within the same time interval from loop detector data. The former is sample since not all vehicles return trajectory data. Pearson correlation coefficient $\rho$ is utilized to evaluate the linear correlation between the link counts extracted from trajectory data $X = \{x_1, \ldots, x_n\}$ and the link counts collected by loop detectors $Y = \{y_1, \ldots, y_n\}$, where $n$ is the size of the statistics, which equals the link number times the interval number.

$$\rho = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$  \hspace{1cm} (14)

$\rho$ ranges from -1 to 1, where $\rho > 0$ is a positive correlation, $\rho < 0$ is a negative correlation, and 0 is no linear correlation. The larger $|\rho|$ indicates a stronger correlation. In general, $|\rho| \geq 0.7$ means a strong linear correlation. Trajectory data are regarded as representative samples if they pass this validation test.

3.4. Weights Calculation. Proper selection of weights is important to the control effect. If the volume reduction of a bottleneck is distributed too much to an on-ramp where vehicles do not destine for the bottleneck, then the metering rate is too restrictive. On the contrary, if too many vehicles are allowed into the freeway but they do not get off the freeway before the bottleneck; the metering rate is too loose. Accurate OD estimation provides valuable information for weights calculation [11]. In this research, the weights calculation is based on the OD patterns extracted from trajectory data. The traffic demand from one on-ramp to different destinations is the on-ramp demand times its destination choices proportion. Trajectory data provide representative samples of OD pairs to calculate this proportion. Let $d_{ij}$ be the traffic demand from on-ramp $j$ to destination $l$, $x_{ij}$ be the corresponding number of trips extracted from trajectory data, and $\nu$ be the number of destinations. Then,

$$d_{ij} = d_j \frac{x_{ij}}{\sum_{l=1}^{\nu} x_{lj}}$$  \hspace{1cm} (15)
Start with the first freeway section $i=1$

Average occupancy > the threshold? (Equation 1)

Section $i$ is not a bottleneck

Go to next freeway section $i=i+1$

Volume entering section > volume exiting section? (Equation 2)

Section $i$ is a bottleneck

Calculate bottleneck volume reduction (Equation 3)

Distribute the volume reduction to upstream on-ramps according to weights (Equation 4)

$N$ has been exhausted?

The most restrictive on-ramp volume reduction is used to calculate the bottleneck metering rate (Equation 5)

Start with the first on-ramp $j=1$

Occupancy of the downstream merging area > the threshold?

Allow the on-ramp volume to make up the difference between downstream capacity and upstream inbound volume

Go to next on-ramp $j=j+1$

Set the local metering rate as the minimum value

Local metering rate (Demand-capacity strategy) (Equation 6)

System metering rate = min (bottleneck metering rate, local metering rate) (Equation 9)

Prevent on-ramp queues exceeding its storage capacity (Equation 10-11)

Adjust ramp metering rate to meet the constraints of queue length and minimum value (Equation 12)

Convert the metering rate to the green time (Equation 13)

Figure 2: The flow chart of the bottleneck algorithm.
The traffic demand from on-ramp $j$ to bottleneck $i$ is calculated by the sum of $d_{ij}$ where destination $l$ locates downstream of bottleneck $i$.

$$d_{ij} = \sum_l d_{ij}$$

(16)

Weights $W_{ij}$ are the weights of on-ramp $j$ to the bottleneck $i$, proportional to traffic demand $d_{ij}$.

$$W_{ij} = \frac{d_{ij}}{\sum_j d_{ij}}$$

(17)

There exists $\sum_j W_{ij} = 1$ for each bottleneck section. Weights matrix is as (18), $u$ and $m$ are, respectively, the number of on-ramps and bottlenecks.

$$W_{u \times m} = \begin{bmatrix} W_{1,1} & \cdots & W_{1,m} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & W_{u,m} \end{bmatrix}$$

(18)

4. Simulation Experiments

4.1. Network Description. A real urban freeway in Shanghai is simulated using the microscopic simulation software PTV VISSIM. The freeway is of 7.5 km in length, with 4 on-ramps, 3 off-ramps, and freeway-to-freeway ramps. The north end locates at a suburban area and the south end is in downtown. Figure 6 annotates the length of mainline sections and on-ramps. The first and second row of the pictures shown in Figure 6, respectively, show on-ramps in the real world and in the simulation.

4.2. Simulation Settings. Experiments simulate freeway operation on weekdays’ morning peak (7:00 – 10:00), a typical recurrent congestion scenario. Data of 24 October 2017 (Tuesday) are utilized to get simulation inputs of traffic demand patterns, as shown in Figure 7. Vehicle Inputs are derived from loop counts. Vehicle routes and their relative flows are set according to the OD patterns extracted from trajectory data. Besides, speed limits on mainlines and on ramps are respectively 80 km/h and 40 km/h, as in the real world.
4.3. Control Parameters Settings. Bottlenecks usually exist at merging areas downstream of on-ramps, weaving areas upstream of off-ramps and lane-drop areas. There are four bottlenecks in this freeway network. Figure 8 shows the relationship between occupancy $o$ and flow $q$ for each bottleneck. A sigmoidal function fits best [40].

$$q(o) = a \cdot c \cdot o \cdot \left(1 - \frac{1}{1 + e^{b - c \cdot o}}\right)$$

where $a, b, c$ are all regression parameters. The occupancy threshold $o_{thres}$ is set corresponding to the capacity flow. When $o_{thres} = [0.28, 0.27, 0.25, 0.22]$, the flow of the bottleneck section reaches the maximum of the fitted curve. In this freeway network, occupancy downstream of the on-ramp and occupancy of the bottleneck section are measured by the same loop detector. Thus, they have the same thresholds. In addition, for all controlled on-ramps, signal cycles $C$ are 60s, saturation flow $r_{sat}$ is 1800 pcu/h·lane, and minimum metering rate $r_{min}$ is 60 pcu/h·lane.

4.4. Scenarios and Cases. The simulation experiments on three cases: (1) no control real-world case; (2) control with loop data—adopting a classic method “gravity model” to estimate OD with loop counts [25]; (3) control with trajectory data—implementing the proposed approach. Control schemes are conducted under three scenarios: (1) real-world traffic demand; (2) 20% less traffic demand for the mainline and all on-ramps; (3) 20% more traffic demand for the mainline and all on-ramps. Besides, to prepare simulation environments, there is no control for the first 15 minutes in all experiments. For each scenario and each case, we run 20 experiments with different random seeds and then take averages.

5. Results and Analysis

5.1. Linear Correlation. There exists a strong linear correlation between link counts from trajectory data ($X$) and those from loop data ($Y$), as shown in Figure 9. Their relation can be fitted as $X = 0.1347Y + 5.815$. This means the penetration of the trajectory data reaches 13.5%, much higher than traditional floating car data. Also, the Pearson correlation coefficient $\rho$ is 0.8775. Thus, the trajectory data pass the validation test and can be used to calculate weights matrix.
5.2. Weights Matrix. There are 4 on-ramps and 4 bottleneck sections, so the weights matrix is in the size of 4×4. Table 1 compares the results of the weights matrix generated from different data sources. Results show that the estimation with only loop data is more easily affected by ramp volumes. Taking the traffic contribution of bottleneck 2 for instance, the result attaches more importance to on-ramp 1 as it has much more volume than on-ramp 2. However, vehicle trajectories reveal that traffic entering from on-ramp 1 massively exits at the immediate downstream off-ramp (off-ramp 1) and does not reach bottleneck 2, because off-ramp 1 connects to another freeway. Similar results can be found in other bottleneck sections. With only fixed-location measurements, the OD estimation deviates more from the reality.

5.3. Mainline Performance. Experiments are performed under three scenarios and three cases. To evaluate the performance, 25 loop detectors are evenly placed on the mainline. The mainline is divided into 25 sections of 300 meters in length on average. Four on-ramps are, respectively, near section 3, 7, 15 and 24. Figure 10 presents mainline occupancies profiles of each experiment, and Table 2 lists performance indicators of average occupancy, average vehicle flow, and mean speed.

Real-world simulation is as in Figure 10(d); the most serious congestion appears in the bottleneck around on-ramp 4. The congestion spreads to upstream mainline over time. Congestion also forms in the rest three bottlenecks sometimes. The ramp metering control with loop data or trajectory data brings significant improvement, as shown in Figures 10(e) and 10(f). Their average occupancies are, respectively, reduced by 18% and 22%. Vehicle flows are increased by 2% and 6%. The congestion is reduced around the fourth bottleneck. It forms later and disappears earlier. Traffic condition is also improved in the rest mainline sections. Congestion is slightly reduced at the beginning and almost eliminated after 75 minutes.

When traffic demand reduces by 20%, traffic congestion only occurs at the fourth bottleneck. To eliminate the congestion, entering traffic is controlled at upstream on-ramps even if there is no congestion around these ramps. Mean speed increases by 5% but vehicle flow decreases by 8%. The control effect with loop data and with trajectory data is of little difference.

When traffic demand increases by 20%, traffic conditions become significantly worse in the first two bottleneck sections. The ramp control with loop data hardly improves the situation. In contrast, the control with trajectory data
Figure 6: Test site with multiple ramps of North-South Elevated Road (southbound) in Shanghai, China.

Table 1: Weights matrix calculated with loop data and trajectory data.

<table>
<thead>
<tr>
<th>Bottleneck 1</th>
<th>Bottleneck 2</th>
<th>Bottleneck 3</th>
<th>Bottleneck 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Loop data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-ramp 1</td>
<td>1.00</td>
<td>0.87</td>
<td>0.38</td>
</tr>
<tr>
<td>On-ramp 2</td>
<td>0.00</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td>On-ramp 3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.56</td>
</tr>
<tr>
<td>On-ramp 4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Trajectory data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-ramp 1</td>
<td>1.00</td>
<td>0.78</td>
<td>0.18</td>
</tr>
<tr>
<td>On-ramp 2</td>
<td>0.00</td>
<td>0.22</td>
<td>0.05</td>
</tr>
<tr>
<td>On-ramp 3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.76</td>
</tr>
<tr>
<td>On-ramp 4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 7: Traffic demand patterns of the test site.

shows better performance, because the control parameters of weights are selected more properly. The control effect is remarkable especially between section 10 and section 25, as more upstream entering traffic can be controlled by ramp metering. It even prevents bottleneck activation around on-ramp 3. Occupancy is reduced by 10% under the control with trajectory data, compared to a 3% reduction under the control with loop data. Vehicle flow and speed are both raised more with trajectory data.

5.4. Overall Travel Time. To reveal whether time saved on the mainline exceeds on-ramp delay caused by ramp metering, the simulation measures travel time of vehicles from different origins, i.e., the upstream mainline and four on-ramps, as shown in Figure 11. For demand level of 100%, the average travel time weighted by flow volume reduces by 19%. The travel time of upstream traffic is reduced by 29%, which means more than 4 minutes are saved for each trip. The travel time of vehicles entering from 4 on-ramps is reduced by 7%, 28%, 23%, and 16%, respectively.

For demand level of 80%, the travel time of traffic from on-ramp 1 and on-ramp 3 is increased by about 75%, which means ramp metering does not bring benefits for these vehicles. However, other traffic’s travel time is saved a little. On average, travel time is increased by 33%. Controls with different data have little difference under these two scenarios.

For demand level of 120%, the average travel time is increased by 6% with loop data and reduced by 11% with
Figure 8: The relationship between occupancy and flow volume at four bottleneck sections.

Figure 9: The linear relationship between link counts from trajectory data and those from loop data.

trajectory data. This result manifests significant advantage of ramp metering with trajectory data. It improves the mainline traffic operation, without the sacrifice of traffic from on-ramps.

6. Sensitivity Analysis

The penetration of trajectory data is supposed to influence the control effect. It is necessary to study its sensitivity on penetration rate and the minimum required value. A simple random sampling is used to lower the penetration rate. In order to reduce random noise, the sampling is done 50 times for each penetration rate and we take averages.

Figure 12(a) depicts the relationship between Pearson correlation coefficient $\rho$ and the penetration rate $p$. With $p$ decreasing, $\rho$ reduces increasingly faster. We need to determine the critical penetration $p_c$ below which $\rho$ drops significantly faster. The second-order difference $\Delta^2 \rho$ is used to evaluate the changing speed of the descending slope. $\Delta^2 \rho$ where $p \leq p_c$ should be significantly larger than $\Delta^2 \rho$ where $p > p_c$. A t-test is applied to determine whether two groups of $\Delta^2 \rho$ divided by $p_c$ are significantly different from each other. When the penetration rate reduces to 4.0%, the t-test rejects the null hypothesis at the default 5% significance level; see Figure 12(b). The penetration of 4.0% is regarded as the minimum value, above which samples of vehicle trajectories are able to represent the characteristics of all vehicles and below which the reliability is not guaranteed.

Simulation experiments are performed under different penetration rates. Figure 13 compares indicators of...
occupancy, vehicle flow, and speed relative to the scenario of no control. Control effects get worse when penetration reduces. When the penetration rate is less than the minimum required value of 4.0%, the performance becomes more sensitive. The reason is that less penetration leads to less reliable results. When trajectory data volume is not sufficient, weight calibration tends to be improper and ramp metering becomes less effective in that way.

7. Conclusion

This paper extracts OD patterns from trajectory data, validates the representativeness of the sample vehicle trajectories, and applies the historical OD patterns to coordinated ramp metering. To verify its advantage, the bottleneck algorithm is implemented with different data sources—loop data (traditional control) and trajectory data (our approach). A real urban freeway is simulated to test control effects.

Simulation experiments show that ramp metering with trajectory data increases another 4% throughput compared to traditional control under current traffic demand. The advantage is more significant under heavier demand. When traditional control can hardly relieve the congestion and even increase average travel time by 6%, our approach increases throughput by 5% and reduces travel time by 11%. However, their control performance shows little difference under low demand. The reason is that trajectory data can help accurate estimation of traffic contribution for bottleneck sections. This information is more important to effective control when traffic condition is oversaturated. Also, the control performance is affected by the penetration of trajectory data. The

<table>
<thead>
<tr>
<th>Demand level</th>
<th>Cases</th>
<th>Occupancy %</th>
<th>% change</th>
<th>Vehicle flow veh/h</th>
<th>% change</th>
<th>Speed km/h</th>
<th>% change</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>No control</td>
<td>21.99</td>
<td>-</td>
<td>2466</td>
<td>-</td>
<td>68.64</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Control with loop data</td>
<td>18.76</td>
<td>-15%</td>
<td>2271</td>
<td>-8%</td>
<td>71.94</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>Control with trajectory data</td>
<td>18.06</td>
<td>-18%</td>
<td>2278</td>
<td>-8%</td>
<td>71.88</td>
<td>5%</td>
</tr>
<tr>
<td>100%</td>
<td>No control</td>
<td>35.74</td>
<td>-</td>
<td>2817</td>
<td>-</td>
<td>57.24</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Control with loop data</td>
<td>29.18</td>
<td>-18%</td>
<td>2883</td>
<td>2%</td>
<td>63.22</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Control with trajectory data</td>
<td>27.90</td>
<td>-22%</td>
<td>2974</td>
<td>6%</td>
<td>65.48</td>
<td>14%</td>
</tr>
<tr>
<td>120%</td>
<td>No control</td>
<td>44.02</td>
<td>-</td>
<td>2616</td>
<td>-</td>
<td>50.01</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Control with loop data</td>
<td>42.55</td>
<td>-3%</td>
<td>2646</td>
<td>1%</td>
<td>50.97</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Control with trajectory data</td>
<td>39.74</td>
<td>-10%</td>
<td>2743</td>
<td>5%</td>
<td>54.02</td>
<td>8%</td>
</tr>
</tbody>
</table>
minimum required penetration is 4.0%, below which the OD estimation tends to become unreliable, and correspondingly the control has no advantage over the traditional one.

The approach proposed in the paper is free of strong assumption and complex computing; hence it is ready for practice. Experiments simulate real road networks and traffic demand patterns. The experiments are more persuasive than those under virtual scenarios. However, there are some limitations of the proposed approach. The bottleneck metering rate is calculated only when bottlenecks have been detected. Also, there exists a time lag between control actions being taken and having effects to the bottlenecks. Thus, it is only suitable to alleviate recurrent freeway congestion with steady OD patterns.

Future researches plan to improve models by further utilizing information recorded by vehicle trajectories. Prediction can be included to activate control before congestion forming. Feedback modules can also be added to increase robustness.

**Data Availability**

The data used to support the findings of this study have not been made available due to data security.
Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by National Natural Science Foundation of China (Grant no. 61773293), Shanghai Sailing Program (Grant no. 19YF1435100), Joint Laboratory for Future Transport and Urban Computing of AutoNavi, and Shanghai Transportation Information Center.

References


