Traffic Information Mining Based on Queue Profile Geometry at Signalized Intersections

Jian Yuan[1,2], Kun An[3], Wanjing Ma[4], Qing Yu[5]

[1]The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai, P.R. China, yuan_jian@tongji.edu.cn
[2]The Department of Civil Engineering, McGill University, Montreal, QC, H3A 0C3, Canada, jian.yuan@mail.mcgill.ca
[3]The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai, P.R. China, kunan@tongji.edu.cn
[4]The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai, P.R. China, mawanjing@tongji.edu.cn
[5]Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology, Shenzhen 518055, China, yuq@sustech.edu.cn

ABSTRACT

Traffic queue profile describes the spatiotemporal formation and dissipation of queues. Existing applications of traffic queue mainly focus on the estimation of traffic state parameters such as traffic volume and queue length. However, the origin of the queuing vehicles from upstream intersections remained unknown unless vehicle identification sensors are installed nearby. In this study, we proposed a traffic demand analysis method based on the geometric features of traffic queue profile. Specifically, starting point, shockwave slope, and queue length are chosen for statistical analysis using a real dataset. This study explores the potential application of traffic queue in the domain of demand analysis, which can be further used in the evaluation and optimization of traffic signal controls.

Keywords: queue profile, geometric analysis, congestion tracing, vehicle trajectories

1. Introduction

Signalized intersections are the main cause of traffic congestion in urban area. While queues at signalized intersections are the most common measurement. To better the dynamics and performance of the traffic system, the problem of how to estimate the queue characteristics has been well studied in the past several decades (Liu et al., 2009; Zhan et al., 2015). Further, the queue profile, which incorporate the time dimension to make a polygon in the time-space plane, aims at modeling the evolution of queues, draws people's attention. It could describe the spatiotemporal formation and dissipation of queues (Ramezani & Geroliminis, 2013). Existing studies employ various monitoring methods, such as video cameras (Puri et al., 2007) and probe vehicle GPS trajectories (Ramezani & Geroliminis, 2013, 2015), to capture the geometry features of the queue profile and achieve satisfied reconstruction performance.

Currently, the application of queue profile mainly focusses on the delay analysis, queue length estimation (Liu et al., 2009), and traffic signal estimation (Wang et al., 2019). Several studies proposed trajectory reconstruction using queue profiles but only limited to continuous flow range (Chen et al., 2021). The full vehicle trajectory that covers different links cannot be reconstructed

 Volume DOI: 10.17605/OSF.IO/N6DU5

due to interruption of traffic signals. The potential information of queue profile geometry has not been sufficiently discussed in the domain of demand analysis. Like a stone falling into the water, the queuing shockwave could be regarded as a reflection of the traffic system on the arriving vehicles. Under a certain scenario, i.e., given the intersection layout, signal schemes, and traffic demand level, the queue profile should also present a certain pattern.

Thanks to the progress in queue profile reconstruction, the queue profile can be extracted in various data conditions. Under this assumption, this paper aims at mining the potential of queue profile in traffic demand analysis. Specifically, formulating the relationship between the shockwave features and the flows from upstream links. Section 2 summarize the method and data, section 3 shows the data investigation results, and section 4 draws the discussion and conclusion.

2. Methodology

To perform traffic demand analysis, especially getting the flow of a certain route across several intersections, estimation of the aggregated link flow is not enough. It is also necessary to trace the proportion of flows come from different upstream links. For example, for a given link a, assuming there are n upstream links. The task is to decompose the link flow y_a into flows from upstream links, i.e., $[y_a^1, \dots, y_a^i, \dots, y_a^n]$. This can be implemented by leveraging the geometry information from queue profile. For example, in Figure 1(1), flows on the selected comes from link 101 and link 102. To fully utilize the geometry information of queue profile, we will first make some field data observations to prove the feasibility of the idea and then illustrate the proposed method.

2.1 Data

In this study, the Next Generation Simulation (NGSIM) dataset is used for modeling and analysis. It is a fully sampled vehicle trajectories data collected by the U.S. department of transportation (Punzo et al., 2011). The vehicle trajectories are extracted through one or several digital video cameras mounted on the top of a building. Specifically, the Lankershim Boulevard dataset, one of several datasets collected under the NGSIM program, is selected for detailed analysis.

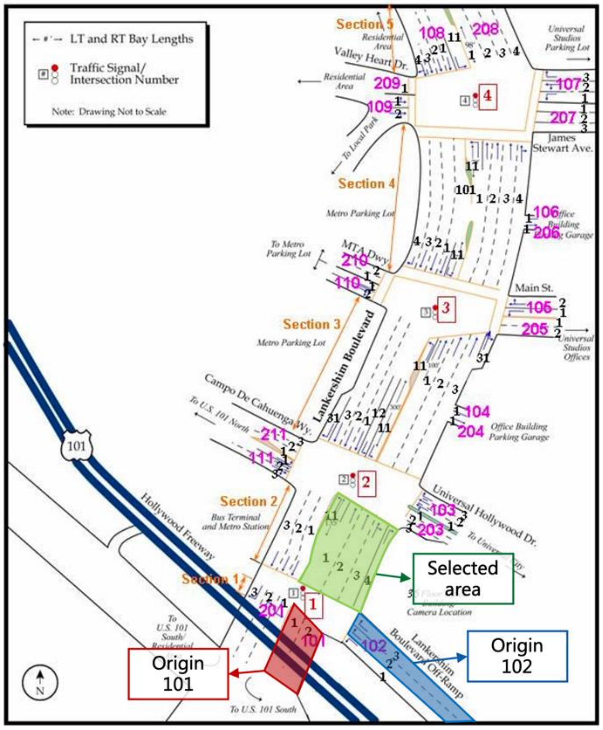
The original dataset includes the vehicle index, location (longitude, latitude, link index, lane index), and speed. Figure 1(1) gives the geometry layout of the study area. As is shown in Figure 1(1), the whole dataset consists of four signalized intersections. Part of the intersection 1 is chosen for further investigation. For three lanes of the selected area (with green mark), the trajectories of upstream flows coming from two different origins, i.e., 101 (with red mark) and 102 (with blue mark), are visualized with different colors.

The dataset is preprocessed through GIS software and Python packages such as TransBigData (Yu & Yuan, 2022) to extract: 1) the traffic signals; 2) the distance between the vehicles and the stop bar; 3) the first stop point of a vehicle in a traffic signal cycle. Then, the trajectories, signal scheme, and the first stop point in a signal cycle are prepared to be visualized in a time-space panel.

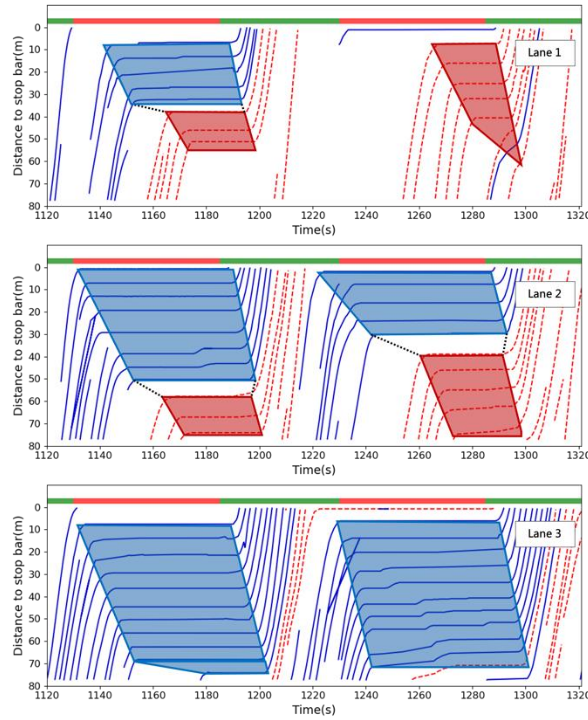
2.2 Methods

In this section, the parameters of the queue profile and the formulation of the flow tracing are proposed.

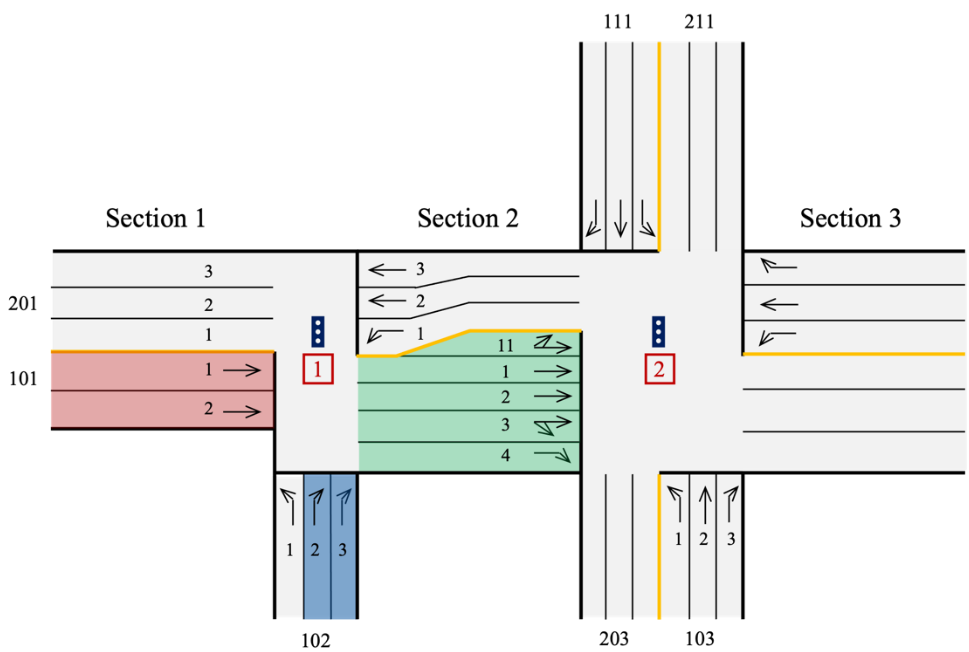
 Volume DOI: 10.17605/OSF.IO/N6DU5



(1) The Selected Area in the Whole Dataset



(2) The Shockwave of Three Lanes



(3) The Geometry of the Selected Area

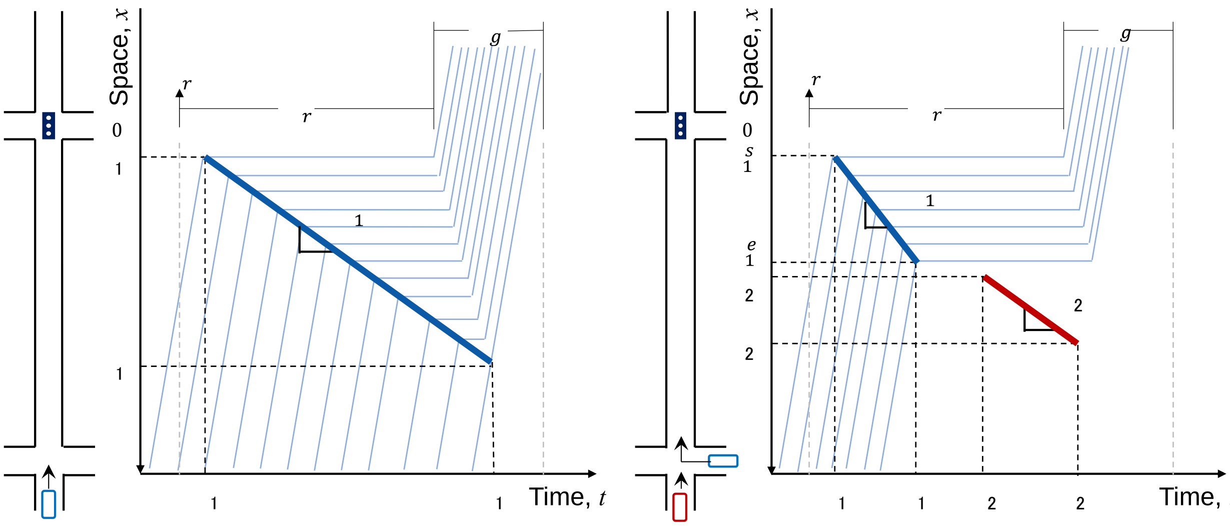
Figure 1 The NGSIM Lankershim Boulevard Dataset

2.2.1 Parameters and feature engineering

(1) Queue profile parameters

Based on above observations, this study proposed the following parameters to characterize the queue profile: 1) the start point (t_1, x_1), 2) the shockwave speed w, and 3) the queue length l. In this study, the fixed signal control plan is considered.

For the traffic signal, denoting t_r as the red start time , d_r as the red duration, d_g as the green duration. For each signal cycle, there may be multi-origin for the queuing vehicles. The scenarios of one-origin and two-origin are provided for illustration in Figure 2(1) and Figure 2(2). The shadow area means the corresponding vehicle speed is zero. The scenario with residual queues (vehicles waited for more than one signal cycle) is not analyzed in this study but provided in appendix.



(1) Single Upstream Source Scenario (2) Two Upstream Sources Scenario

Figure 2 The Basic Parameters of Queue Profile as Signalized Intersection

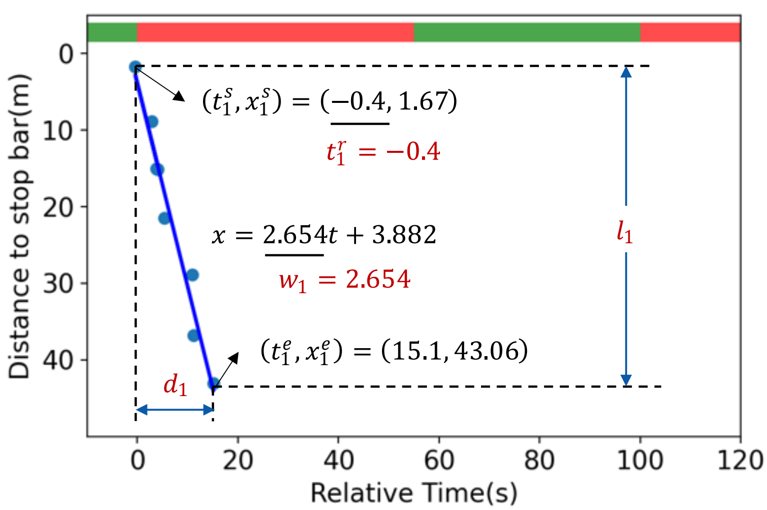


Figure 3 Examples of Features Extracted from Queue Profile

(2) Feature engineering

Feature engineering is crucial for better capturing the relationship between flow and queue profile geometry, for example, using statistical or artificial intelligence methods. Assuming several segmented shockwaves, $1, \dots, i, \dots m$, are well observed in a single traffic signal. For each shockwave segment i, the start/end location/time is denoted as (x_i^s, t_i^s) and (x_i^e, t_i^e). A real data case from NGSIM dataset is provided in Figure 3. Then, the following features can be extracted:

Table 1 Queue Profile Features

Features of the ith Queue Profile	Expressions
Start location	x_i^s
Relative start time in the signal cycle	$t_i^r = t_i^s - t_r$
Maximum length	$l_i = x_i^e - x_i^s$
Duration time	$d_i = t_i^e - t_i^s$
Profile slope	$w_i = \dfrac{x_i^e - x_i^s}{t_i^e - t_i^s}$

In this study, the above features of queue profile in the NGSIM dataset are extracted. Firstly, a stop detection method is developed (see Figure 3 for the detection results). Then, for each lane and each traffic signal cycle, the queue profile parameters, and its label (true origin) are extracted.

2.2.2 Formulation of the flow tracing model

The framework consists of two part: 1) matching the queue profile to the corresponding origins, and 2) estimating the corresponding traffic volume. Since the traffic volume can be estimated using the queue length and the average headway, this study only focuses on the flow tracing method.

As stated in the literature review, we assume that the queue profile features are already known. The flow tracing task can be treated as a classification problem using logistic regression, a type of supervised machine learning method to predict one of two possible outcomes, often represented as 0 and 1.

In this case, the model inputs include start location, relative start time in the signal cycle, maximum length, and profile slope. Duration time is not included since the duration time d_i can be expressed by the queue length l_i and profile slope w_i. For the ith shockwave, the simple formulation is as follows:

$$z = \mathbf{f}[x_i^s, t_i^r, l_i, w_i] \tag{1}$$

where $\mathbf{f}$ represents the logistic regression method, z is a binary variable which represent the origin of the shockwave. In this case, two origin flow sources are link 101 and link 102 (Figure 1).

This model can be extent to multi-classification problem if there are more than two upstream links. All input features are normalized before the training process.

3. Results

3.1 Data extraction and analysis

In this section, the NGSIM vehicle trajectories is used to investigated the proposed method. Firstly, the stop information is extracted. The queue profiles of three lanes (1, 2, 3) of selected area (shown in Figure 1), among 20 signal cycles, were extracted for model training and analysis. Figure 4**Error! Reference source not found.** shows the trajectories collected during 1000~2000 seconds and the stop locations.

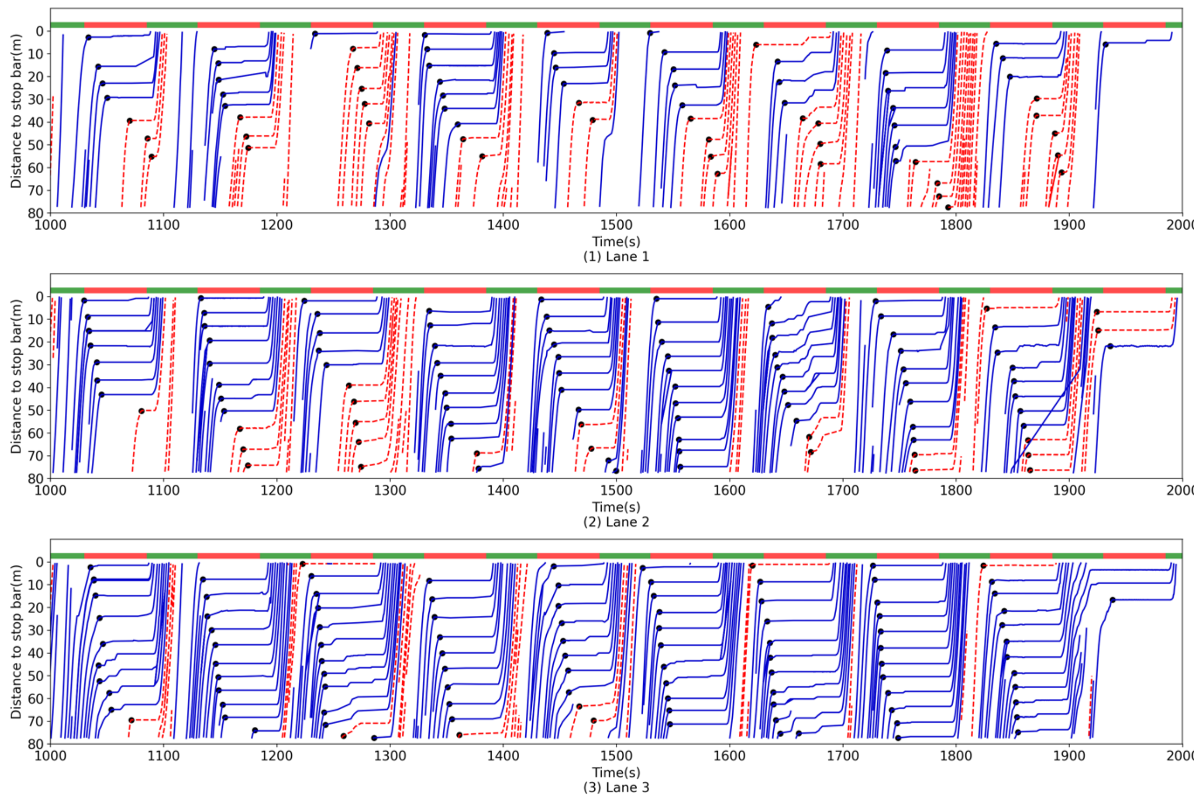


Figure 4 Stop Points Recognition of Three Lanes in the Selected Area

Figure 5 shows the first stop time-space information over several signalized cycles. The results shows that the time-space locations from different upstream links also differ. Vehicles from link 101 are more likely to stop in two time-space areas:

- During transition from green to red, close to the stop bar
- From the middle to the end of the red light, farther away from the stop bar

Thus, the statistical probability of vehicle first stop location should be different among vehicles from different locations. In addition, both Figure 1(2) and Figure 4 **Error! Reference source not found.**shows that the length of the queue profile could reflect the number of vehicles. This is intuitive since vehicles tend to keep a proper headway. Thus, the spatial length, i.e., l_i, should

 Volume DOI: 10.17605/OSF.IO/N6DU5

also be an informative input for the model. These finding proves the rationality of the aforementioned model.

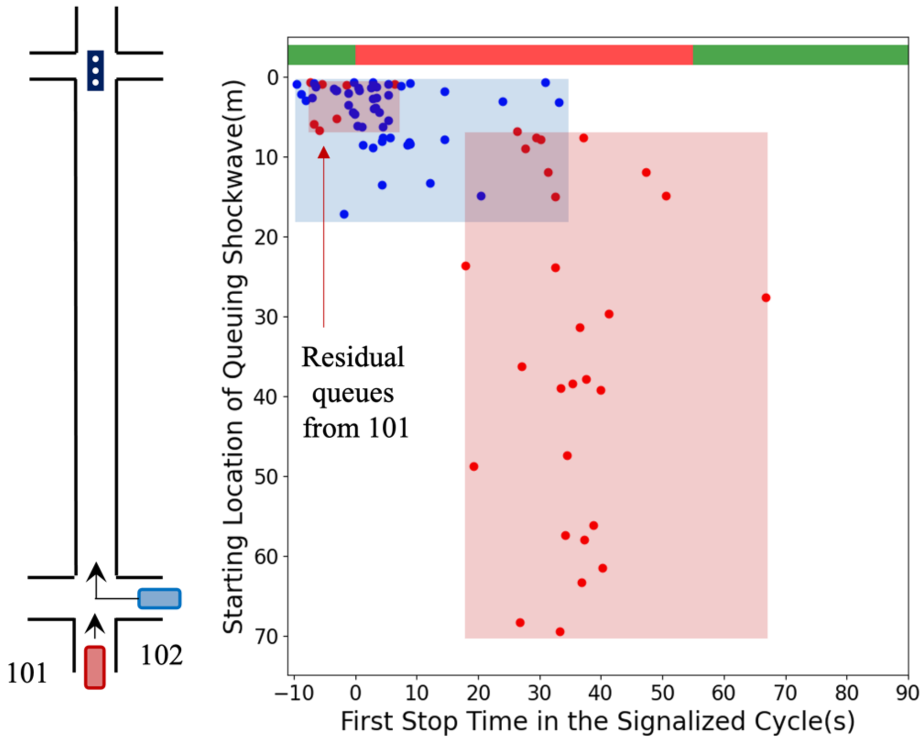


Figure 5 The Distribution of the Queue Profile Starting Points

3.2 Model performance

In the logistic regression model, 80% data is used for training and the rest for testing. The detailed information of each lane and the model performance is shown in Table 2. The matching accuracy of lane 1 and lane 2 is around 90% while lane 3 shows a lower accuracy. The lack of sufficient training data may account for the various prediction performance among different lanes. Overall, the results show that it can be used to analysis the flow volume composition.

Table 2 Prediction Accuracy of the Proposed Method

Lane ID	Number of recorded queue profiles		Prediction Accuracy
	From link 101	From link 102	
Lane 1	20	25	88.89%
Lane 2	21	26	90.00%
Lane 3	16	29	66.67%

4. Discussion & Conclusion

Queue profile at intersections could reflect the dynamics of the urban traffic system. Despite various methods for estimating the queue profile, current applications mainly focus on traffic performance analysis rather than demand analysis. In this study, the queue profile of NGSIM dataset is visualized and differentiated with different origins. The stratification phenomenon and the queue profile spatiotemporal location is found to be informative for tracing the downstream link flows. Based on field data observations, a logistic regression model is proposed for tracing the link flow to upstream flows, which utilize features like queue profile start location, start time, slope, and maximum length. The results show that the method could be successfully applied in the flow-tracing task, thus help improve the traffic control strategies.

This study provides some new insights of queue profile in the application aspect and could inspire more topics. In the future, simulation experiments can be implemented to investigate the mechanism between queue profile and traffic flows. For example, the impact of various traffic signal control parameters, demand levels, and different lane changing behaviors. In addition, methods are required to automatically recognize the correct origins of the segmented queue profile. The queue profile reconstruction quality should also be considered in building real applications.

Acknowledgments

This research is sponsored by the China Scholarship Council (202106260122).

Appendix

Figure 6 shows the queue profile when there are residual queues which cause second stop. It reflects that the signal settings may affect the geometry of queue profiles.

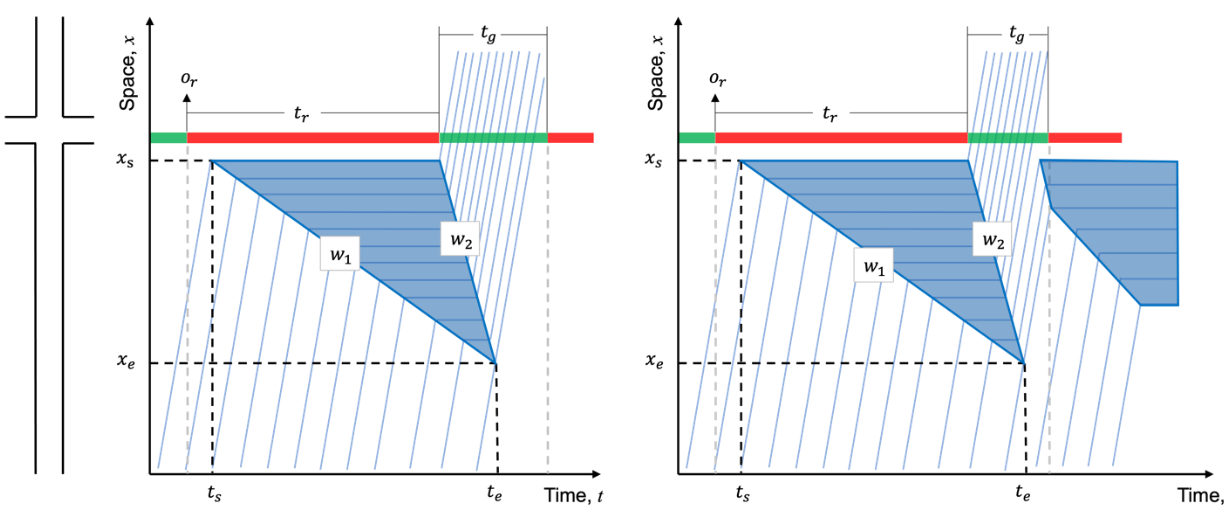


Figure 6 Queue Profile with Residual Queues

Volume DOI: 10.17605/OSF.IO/N6DU5

References

Chen, P., Wei, L., Meng, F., & Zheng, N. (2021). Vehicle trajectory reconstruction for signalized intersections: A hybrid approach integrating Kalman Filtering and variational theory. *Transportmetrica B: Transport Dynamics*, *9*(1), 22–41. https://doi.org/10.1080/21680566.2020.1781707

Liu, H. X., Wu, X., Ma, W., & Hu, H. (2009). Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies*, *17*(4), 412–427. https://doi.org/10.1016/j.trc.2009.02.003

Punzo, V., Borzacchiello, M. T., & Ciuffo, B. (2011). On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. *Transportation Research Part C: Emerging Technologies*, *19*(6), 1243–1262. https://doi.org/10.1016/j.trc.2010.12.007

Puri, A., Valavanis, K., & Kontitsis, M. (2007). Generating Traffic Statistical Profiles Using Unmanned Helicopter-Based Video Data. *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 870–876. https://doi.org/10.1109/ROBOT.2007.363095

Ramezani, M., & Geroliminis, N. (2013). Exploiting probe data to estimate the queue profile in urban networks. *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 1817–1822. https://doi.org/10.1109/ITSC.2013.6728492

Ramezani, M., & Geroliminis, N. (2015). Queue Profile Estimation in Congested Urban Networks with Probe Data. *Computer-Aided Civil and Infrastructure Engineering*, *30*(6), 414–432. https://doi.org/10.1111/mice.12095

Wang, S., Huang, W., & Lo, H. K. (2019). Traffic parameters estimation for signalized intersections based on combined shockwave analysis and Bayesian Network. *Transportation Research Part C: Emerging Technologies*, *104*, 22–37. https://doi.org/10.1016/j.trc.2019.04.023

Yu, Q., & Yuan, J. (2022). TransBigData: A Python package for transportation spatio-temporal big data processing, analysis and visualization. *Journal of Open Source Software*, *7*(71), 4021. https://doi.org/10.21105/joss.04021

Zhan, X., Li, R., & Ukkusuri, S. V. (2015). Lane-based real-time queue length estimation using license plate recognition data. *Transportation Research Part C: Emerging Technologies*, *57*, 85–102. https://doi.org/10.1016/j.trc.2015.06.001