

Research Article

Modeling and Predicting Stochastic Merging Behaviors at Freeway On-Ramp Bottlenecks

Jian Sun ¹, Kang Zuo,² Shun Jiang,¹ and Zuduo Zheng³

¹Department of Traffic Engineering & Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, China

²Wuhan Planning & Design Institute, 1250 Jinghan Road, Wuhan, China

³School of Civil Engineering, The University of Queensland, St. Lucia 4072, Brisbane, Australia

Correspondence should be addressed to Jian Sun; sunjian@tongji.edu.cn

Received 12 December 2017; Revised 26 March 2018; Accepted 15 April 2018; Published 16 May 2018

Academic Editor: Emanuele Crisostomi

Copyright © 2018 Jian Sun 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.

Merging behavior is inevitable at on-ramp bottlenecks and is a significant factor in triggering traffic breakdown. In modeling merging behaviors, the gap acceptance theory is generally used. Gap acceptance theory holds that when a gap is larger than the critical gap, the vehicle will merge into the mainline. In this study, however, analyses not only focus on the accepted gaps, but also take the rejected gaps into account, and the impact on merging behavior with multi-rejected (more than once rejecting behavior) gaps was investigated; it shows that the multi-rejected gaps have a great influence on the estimation of critical gap and merging prediction. Two empirical trajectory data sets were collected and analyzed: one at Yan'an Expressway in Shanghai, China, and the other at Highway 101 in Los Angeles, USA. The study made three main contributions. First, it gives the quantitative measurement of the rejected gap which is also a detailed description of non-merging event and investigated the characteristics of the multi-rejected gaps; second, taking the multi-rejected gaps into consideration, it further expanded the concept of the “critical gap” which can be a statistic one and the distribution function of merging probability with respect to such gaps was analyzed by means of survival analysis. This way could make the full use of multi-rejected gaps and accepted gaps and reduce the sample bias, thus estimating the critical gap accurately; finally, considering multi-rejected gaps, it created logistic regression models to predict merging behavior. These models were tested using field data, and satisfactory performances were obtained.

1. Introduction

Numerous studies have shown that merging from the acceleration lane has a significant impact on traffic operations at freeway on-ramp bottlenecks [1, 2] and can also trigger traffic breakdown [3–5]. Merging behavior is a complex task, because, typically, the driver has to focus on three or more vehicles in their current and target lanes in a limited timeframe [5]. This contrasts with the driver's task in the car-following process where she/he generally only needs to focus on the immediately preceding vehicle in their current lane [6]. Therefore, the study of merging behaviors is a challenging one.

In recent decades, merging behavior has been extensively studied, and many models have been developed. Lots of these models are based on the gap acceptance theory; that is, when

meeting a gap, a driver will compare it to the critical gap. If the gap is larger than the critical gap, the driver will accept it and merge; if not, he/she will reject it and move on to find another gap [7, 8]. Intuitively, there are two types of gap: the accepted gap and the rejected gap. Many previous studies are related to the former, while few concern the latter. In addition, to the best knowledge of the authors, no study has investigated multi-rejected gaps and their impact on merging behavior.

This study aimed to fill this gap in the literature by focusing on the relationship between multi-rejected gaps and merging behavior. In so doing, it made three main contributions: first, it redefines the rejected gap and gives the quantitative measurement of it and investigated the characteristics of the multi-rejected gaps in Los Angeles and Shanghai, respectively; second, it extended the concept of the “critical gap” that could be a stochastic one when

taking into account the multi-rejected gaps and used survival analysis to estimate the merging probability function with respect to such gaps; finally, considering multi-rejected gaps, logistic regression models are created to predict merging behaviors. These models have been tested using field data, and satisfactory performances have been obtained. With the proposed model, the microscopic simulation is able to reproduce traffic operation more accurately. In future, the autonomous vehicles will benefit from the model when predicting surrounding vehicles near on-ramps and merging reasonably.

The remainder of this paper is structured as follows: Section 2 presents a review of the literature related to merging behavior; Section 3 describes the study sites and data; Section 4 discusses characteristics of the multi-rejected gaps; Section 5 expands the concept of the critical gap which can be a stochastic one and its estimation method; Section 6 describes the logistic models to predict merging behavior with the consideration of multi-rejected gaps; and Section 7 concludes the paper by summarizing its main findings and recommending topics for future study.

2. Literature Review

As demonstrated above, this study aimed to analyze the relationship between multi-rejected gaps and merging behavior. This section discusses the existing literature on the three specific aspects of this relationship which need to be addressed.

The first task is to gain an understanding of the concept of a “rejected gap.” Initially, it is used in the estimation of critical gaps at unsignalized intersections [9]. However, little attention has been paid to the rejected gap in a merging area, and most scholars simply regard rejected gaps as non-merging events. Hou et al. [10] note that when a vehicle’s lateral coordinates remain the same lane or with some oscillations, it is a non-merging event. Similarly, Meng and Weng [11] define a rejected gap as the situation in which the driver fails to merge into the current adjacent lane. Daamen et al. [12] note that drivers prefer to choose an optimum gap that might reject several acceptable gaps before merging. Marczak et al. [7] first define the rejected gap as the gap a merger could have chosen (but chooses instead to drive ahead and merge into a gap downstream) and analyze the impact of one rejected gap. Based on the evidence of these studies, therefore, we conclude that previous studies either lack a clear description of rejected gaps or simply consider one rejected gap. Furthermore, the selection and calculation of rejected samples in these studies are also not clearly defined.

The second aspect of the relationship between multi-rejected gaps and merging behavior requiring analysis relates to gap acceptance theory [13–15]. This theory is extensively used in the modeling of merging behaviors, and the critical gap is its key. The theory is mostly based on the assumption that rejected gap \leq critical gap \leq accepted gap [16]. From this equation, we can easily see that the accepted gap is no less than the rejected gaps. However, according to our research, this is not often the case: the rejected gap can be larger than the accepted gap (see Section 4.2 for details).

Gap acceptance theory is also widely used in microscopic traffic simulation models. Current simulation models, such as MITSIM [17], CORSIM [18], VISSIM [19], and TransModeler [20], use critical gaps in different ways. For example, risk factors are used to define the critical gaps in CORSIM; and a psychophysical model is used in VISSIM to obtain a critical gap. The TransModeler defines the linear and non-linear critical gaps according to a combination of the speed of the subject vehicle, the lead gap, and the lag gap in the target lane. However, these simulation models fail to realize the stochastic nature of the merging behavior. In fact, even on the basis of gap acceptance theory, merging is a stochastic rather than deterministic event. In addition, the critical gap is also stochastic rather than constant. A lognormal distribution of critical gap is assumed in MITSIM. Although it notices the stochastic nature of critical gap, it neglects the influence of the rejected gap on the critical gap. Hence, the assumed distribution of the critical gap is unsuitable [16].

Finally, an accurate prediction of merging behavior or decisions is the basis of traffic simulation and driver assistance systems. This is not only because merging behavior might lead to traffic breakdown, but also because a last minute merging decision will affect subsequent driving behavior and the behaviors of the following drivers [21]. Hidas [22, 23] developed a merging model including cooperative and forced merge movement components. Most of existing merging models considered the effects of the instant speed, the relative speed, and the gaps of merging vehicle with its assumptive lead and lag vehicle [24, 25]. With further study, more surrounding traffic characteristics were taken into consideration for modeling merge behaviors. Hidas [22, 23] considered the effect of the reaction time, the maximum waiting time, and time distance on merging behaviors. Sun and Elefteriadou [26] modeled this behavior considering driver characteristics. Recently, Marczak et al. [7] take the once rejected gap before merging into consideration to model the merging behavior.

As for the approach to model the merging behavior, it can be either parametric or non-parametric. Discrete choice models such as the binary logit model, the multinomial logit model, and the nested logit model are proposed to model lane-changing behavior [27]. The models mentioned above are parametric. In recent years, however, non-parametric machine learning methods such as decision tree [11], fuzzy genetic algorithm [10], and Bayes classification [28] have also been introduced to model lane-changing behavior. However, in their application of these models, all the above-mentioned studies neglected the influence of multi-rejected gaps on the merging prediction.

In summary, studies reported in the review of the literature (1) do not clearly define and calculate the rejected gap; (2) ignore the critical gap which can be a stochastic one (caused by rejected gaps) and the real relationship between the accepted and rejected gap; and (3) fail to consider multi-rejected gaps in most gap acceptance model and merging prediction models. These deficiencies are addressed in this study.

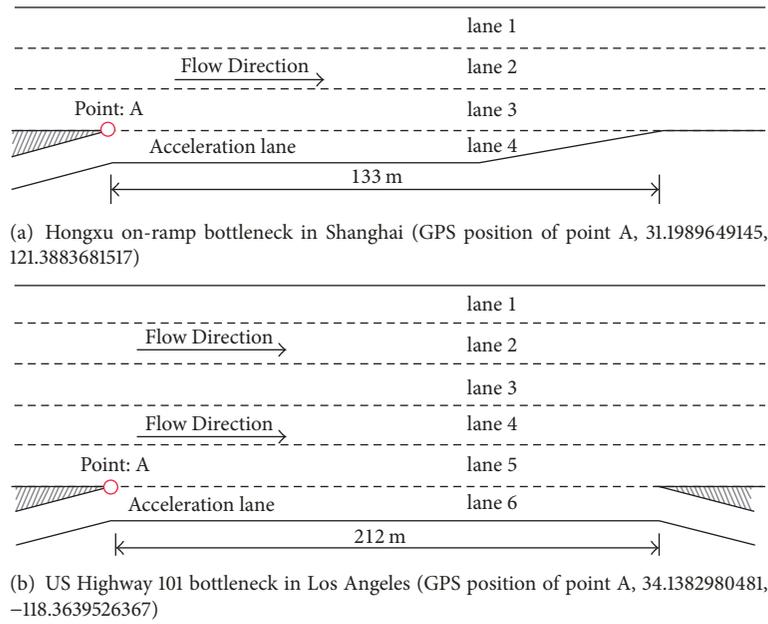


FIGURE 1: The schematics of two study sites.

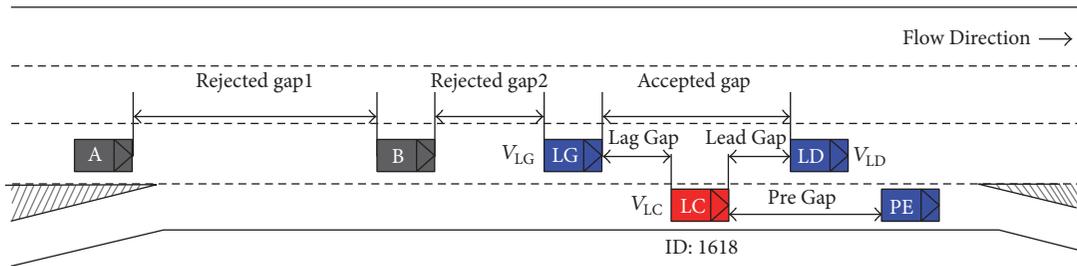


FIGURE 2: Description of the variables in a merging process, where LC is the merging vehicle, LD is the leading vehicle on the target lane, LG is the following vehicle on the target lane, and PE is the leading vehicle on the initial lane.

3. Study Sites and Data

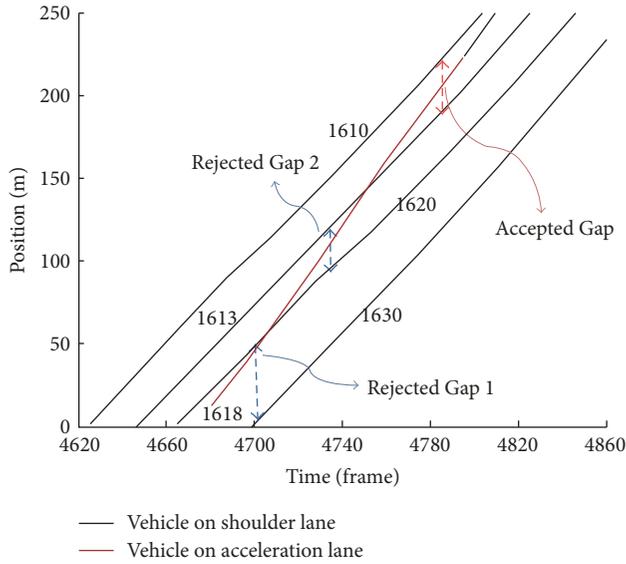
Aimed to better understand the merging behavior affected by multi-rejected gaps, two isolated bottlenecks were picked, and US-101 NGSIM data set and Hongxu on-ramp data set were selected to ensure sufficient rejected gaps and a relatively higher data resolution.

3.1. Study Sites. Figure 1 shows the schematic diagrams of the two on-ramp bottleneck sites. Figure 1(a) illustrates the Hongxu on-ramp bottleneck, on an eastbound section of the Yan'an Expressway in Shanghai (China), hereafter referred to as "SH." There are three mainline lanes, and the acceleration lane has a length of 133 m. Figure 1(b) illustrates a southbound segment of the US-101 in Los Angeles (USA), hereafter referred to as "LA." In this case, there are five mainline lanes, and the acceleration lane has a length of 212 m. During the observation time, both sites are in congested condition, and the speed is about 10 km/h to 65 km/h in SH and 30 km/h to 50 km/h LA, respectively. It should be pointed out that there are also diverging vehicles in LA; however, it is quite less than

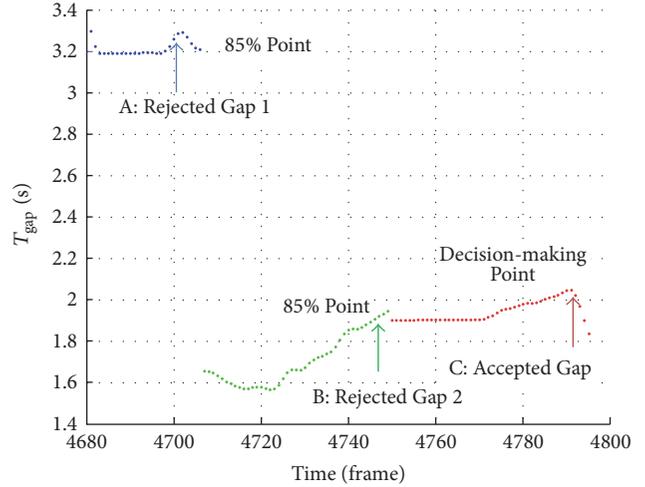
the merging vehicles and we remove the merging events that are obviously affected by the diverging vehicles.

3.2. Data Extraction. To collect data from SH, video cameras were installed in buildings 25 m or more tall. The cameras covered 80 m~100 m upstream and approximately 20 m downstream of the bottleneck, where merging behaviors are prevalent. Trajectories were extracted with the help of advanced track processing software-George (also see in [8]). This software can manually identify and mark the initial position of each vehicle in the video and then track its trajectory. It can record vehicle position, velocity, acceleration, and other parameters at 0.12 s intervals [8]. The data for LA are from the NGSIM dataset, with a time resolution of 0.1 s. These LA data collection details can be found in Zheng et al. [29].

The study focuses on merging vehicles and the vehicles surrounding them. 332 and 609 merging samples were collected from SH and LA, respectively. Similar to Sun et al. [27] and Marczak et al. [7], variables considered in the analysis are depicted in Figure 2 (detailed trajectory is show in Figure 3) and summarized in Table 1. For clarity, the meanings



(a) The trajectories of merging Vehicle 1618 and its surrounding vehicles (the red line represents the vehicle on the acceleration lane, while the black line represents the vehicle on the shoulder lane)



(b) The calculation of Vehicle 1618's rejected gaps and accepted gap

FIGURE 3: An example of a merging process at LA.

TABLE 1: Potential variables in the merging process.

Indicator	Notation	Unit
Remaining distance on acceleration lane	D	m
Speed of merging vehicle	V	m/s
Time gap*		
LD and LG	T_{gap}	
LD and LC	T_{lead}	s
LC and LG	T_{lag}	
PE and LC	T_{pre}	
Space gap		
LD and LG	S_{gap}	
LD and LC	S_{lead}	m
LC and LG	S_{lag}	
PE and LC	S_{pre}	
Speed difference		
LD and LC	ΔV_{lead}	m/s
LC and LG	ΔV_{lag}	

*Throughout the article, "gap" indicates time gap unless explicit noted.

of variables are described in Figure 2 taking Vehicle 1618 in LA as an example.

$$\begin{aligned} \Delta V_{\text{lead}} &= V_{\text{LD}} - V_{\text{LC}} \\ \Delta V_{\text{lag}} &= V_{\text{LC}} - V_{\text{LG}} \end{aligned} \quad (1)$$

where V_{LD} is the speed of leading vehicle at the shoulder lane; V_{LC} is the speed of merging vehicle; V_{LG} is the speed of following vehicle at the shoulder lane.

3.3. Defining and Measuring the Rejected Gaps. Previous studies give the various descriptions of the rejected gap. In this study, we used the concept of rejected gap as Marczak et al. [7] mentioned that the rejected gap is the gap a merger could have chosen. However, the selection and calculation of rejected samples and their gaps are not clearly explained in all the previous studies. This study addresses this shortcoming, and the detailed process is shown below. The vehicle trajectories, the rejected gaps, and the accepted gap for the merging process of Vehicle 1618 in LA dataset are shown in Figure 3. The merging vehicle 1618 is faster than other vehicles. Vehicle 1618 experienced two rejected gaps while driving on the acceleration lane (the red trajectory). Then it accepted the gap between Vehicle 1610 and Vehicle 1613 and merged into the mainline (the black trajectory of 1618). Points A and B denote two rejected gaps, while point C represents the accepted gap (see Figure 3(b)).

In Figure 3(a), the intersection of the red and black lines means that the merging vehicle overtakes the vehicle in the mainline, rather than merging. That is, before this point, the merging vehicle rejected a gap. The point where a line changes from red to black is the merging point (when the midpoint of a vehicle's front bumper crosses the lane marking, the corresponding time point is the point at which that vehicle changes lane). In this instance, Vehicle 1618 entered the observation area of the mainline at the 4680 timeframe. Before it merged into the mainline at the 4791 timeframe (accepting the gap between Vehicles 1610 and 1613), it has rejected two gaps: one between Vehicles 1620 and 1630, and the other between Vehicles 1613 and 1620.

Similarly, for each pair of lead and lag vehicles at the shoulder lane, we can calculate the gap at each time step (e.g., 0.1 s for LA data). Thus, a merging vehicle experiences a series

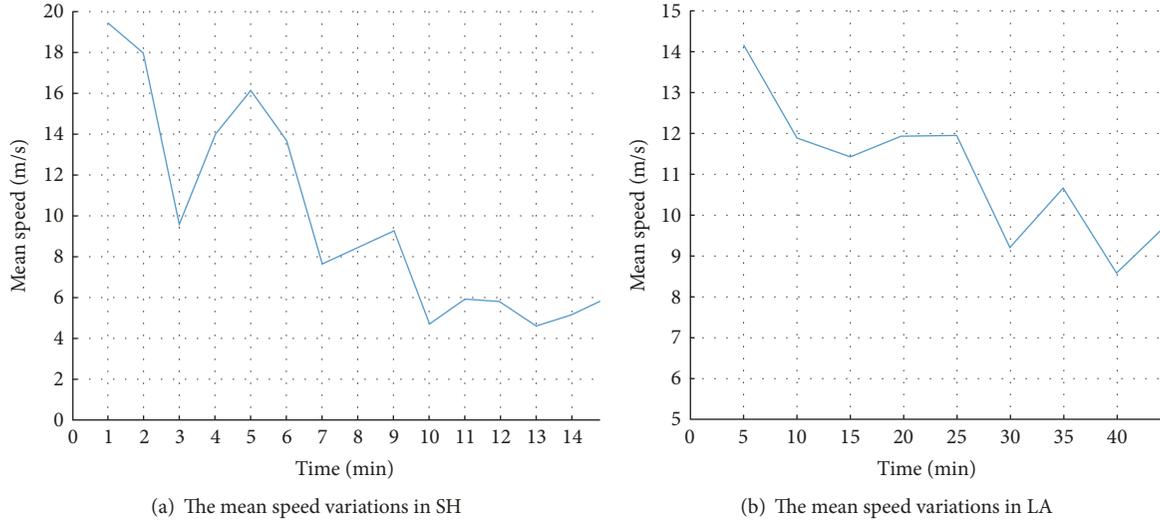


FIGURE 4: The mean speed variations in two datasets.

TABLE 2: Statistics of traffic flow states.

Dataset	Duration time (min)	Range of Mean speed (m/s)	Range of Mean Flow (veh/h/ln)
SH	15	4.4~19.2	1500~1812
LA	45	8.6~14.1	1450~1840

of gaps before the merge is eventually executed. The accepted gap in this study is selected and calculated at 0.5 s before the merging point, but which gap should be considered as the rejected gap for this merging vehicle? There are two steps to calculate and select the rejected gaps.

Step 1. Calculate the 85th percentile as the rejected gap (T_{gap}). It is for the following two reasons: (1) the 85th percentile is often used in traffic engineering; for example, the speed limit is usually set as the 85th percentile [30]; (2) the 85th percentile of the gaps is large enough for drivers to perceive and to make merging decision accordingly. Meanwhile, we also tested the 15th and 50th percentiles of the rejected gaps and found no significant difference among them. We have made the mean test and distribution test for the each other of those three percentiles. It shows that all the P value is higher than 0.05; therefore there are no significant differences between two test object no matter for mean or distribution.

Step 2. Exclude the unreasonable gaps. There are several standards to do the data cleaning: (1) the mean (E) and standard deviation (σ) of merging samples T_{gap} on both sites are calculated, respectively, and the rejection samples whose T_{gap} is smaller than $(E - \sigma)$ are removed; (2) the gaps that drivers do not get prepared to merge in are eliminated; and (3) the samples that the speed of merging vehicles is too slow (e.g., 10 km/h) to execute a lane change process are also excluded from the rejected gaps. We believe that these two steps ensure that the selected gaps are reasonable.

4. Statistical Analysis of Rejected Gaps during the Merging Process

4.1. Statistics of Traffic Flow States. The states of traffic flow in SH and LA datasets are shown in Figure 4 and Table 2. The speed, respectively, range from 4.4 m/s to 19.2 m/s in SH and 8.6 m/s to 14.1 m/s in LA, and during most of the observation time, the speed is a relative low one. Therefore, the proposed model is suitable for the merging behavior prediction in middle-and-high-density flow. The key property in this study is that the merging speed is larger than the mainline speed. In this condition, merging vehicles are able to overtake the preceding mainline vehicles and select a suitable gap from several ones. Therefore, the rejected gaps have significant impact on merging decision.

4.2. Distribution of Rejection Number. The rejection number is the number of rejected gaps before a successful merge. Its distribution in SH and LA is shown in Figure 4.

As shown in Figure 5, as the rejection number increases, the number of vehicles decreases. Indeed, more than 74.7% of vehicles in SH and 70.1% of vehicles in LA rejected a gap at least once. The mean rejection number is 2.2 in SH and 2.0 in LA.

4.3. Relationship between Rejected Gaps and Accepted Gaps. As previously discussed, there might be several rejected gaps during a vehicle's merging process. The following analysis now determines whether the rejected gaps can be larger than the accepted gap during one merging process.

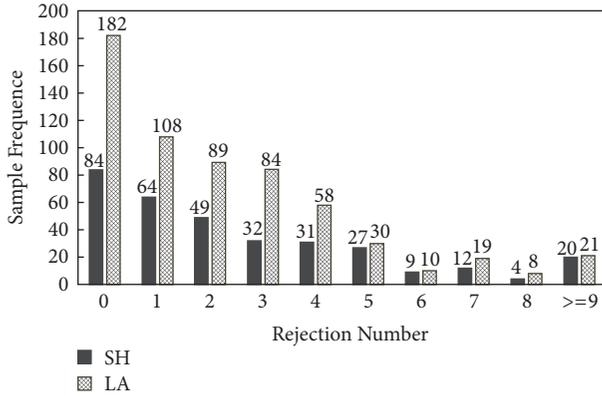


FIGURE 5: The distribution of the rejection number.

TABLE 3: Percentage of merging event that at least one rejected gap larger than accepted gap for samples with different rejection number.

Rejection number	1	2	3	4	5
SH	56.2%	71.4%	62.5%	64.5%	92.6%
LA	55.0%	69.7%	78.0%	80.0%	69.0%

First, the distributions of S_{gap} and T_{gap} for both accepted gaps and rejected gaps in SH and LA are shown in Figure 5. This figure clearly reveals a large overlap for either S_{gap} or T_{gap} at both sites, thus indicating that the rejected gaps can be larger than the accepted gap. Moreover, using t -test, rejected and accepted gaps are statistically compared. Results show that the mean of accepted S_{gap} is significantly larger than the mean of rejected S_{gap} both in SH and LA, while the difference is 9.3 m in LA and 5.9 m in SH, respectively. When it comes to T_{gap} , the mean of accepted T_{gap} is not significantly different to the rejected T_{gap} in SH and LA. And P value in SH ($P = 0.25$) is larger than that in LA ($P = 0.19$). One possible explanation is that drivers in SH are more irrational and thus engage in more aggressive merging behaviors (e.g., forced merging), as reported in Sun et al. [4, 27].

To further verify the findings above, the percentage of the merging events with at least one rejected gap larger than the accepted gap for samples with different rejection number (i.e., rejection number > 0) are listed in Table 3. Similar results (rejected gaps can be larger than the accepted gap) are obtained. In particular, Table 3 shows that the percentage is relatively large (reach 92.6%) for SH. This finding indicates that SH drivers are more unpredictable (an indirect indication of aggressiveness), while LA drivers are more rational and observant (Figure 6).

4.4. Merging Features for Different Rejection Numbers. This section focuses on the merging features related to different rejection numbers.

As for the merging behavior of individual vehicle, we compared the speed, speed differences, space gaps, and time gaps of rejected gaps and accepted gaps of each vehicle. However, because of the driver heterogeneity, there is no obvious conclusion obtained. This aspect is the key point in our further research.

As for the aggregate merging behavior of all vehicles, in accordance with the studies of Marczak et al. [7] and Sun et al. [4], variables that have a significant impact on merging behavior are analyzed, including T_{gap} , D , V , ΔV_{lag} , and S_{gap} . Samples with the same rejection number are put together and the accepted value is analyzed. The box-plots of these merging features in relation to different rejection numbers are shown in Figure 7.

Figure 7 shows that, in both SH and LA, the critical T_{gap} (means of accepted gap) is relatively stable (i.e., mostly 2~3 s). This finding not only emphasizes the key critical gap, but also suggests the appropriate time gap for a merging decision.

As for D , we can see that, with the increase in rejection number, the merging location is closer to the end of the acceleration lane and the distribution is more concentrate for both SH and LA. This implies that the more the gaps are rejected, the less the space can be chosen and the more urgent the merging event becomes. This phenomenon may cause a higher probability of executing a forced lane changing and donates great damage to the bottleneck, which accords with what Sun et al. [4] finds.

V in relation to different rejection numbers differ from SH to LA. In SH, the merging speed decreases with the increase of rejection number, which is opposite to LA. It because that the speed in the acceleration lane is faster than that in the shoulder lane in SH, and drivers reject more gaps to reduce the time cost, while LA drivers reject several gaps to seek a better lane change condition.

Figure 7 shows that with the increase of rejection numbers the speed difference between LC and LG in SH is always fluctuating around 0. The small speed difference and time gap demonstrate the irrational merging behavior, which may cause more cooperative and forced lane changings. In this way, the vehicle in the target lane must slow down actively or passively. According to research finds of Sun et al. [4], there are more forced lane changings in SH. This implies that merging decisions in SH are more aggressive and selfish. However, the speed difference between LC and LG increases with the increase of rejection numbers in LA (over 6 m/s), which implies that merging decisions in LA are more rational.

Figure 7 show a diversity merging behavior under different rejection number. It means that rejected gaps have impact on merging behaviors. These figures also show a different merging preference between drivers in SH and LA. Specifically, drivers in SH are more risk-taking and self-focused while drivers in LA are more rational and altruistic. This results in more forced and cooperative lane changings in SH [27] and more free lane changings in LA.

5. Estimation of Critical Gap and the Distribution Function

As discussed above, during the merging process, a vehicle might reject several gaps, and the rejected gaps might be larger than the accepted one. Therefore, it is reasonable to treat the gap acceptance process of a merging event as stochastic, rather than deterministic. In other words, whether drivers choose to merge or not under a given gap can be characterized as a probabilistic event. Correspondingly, the

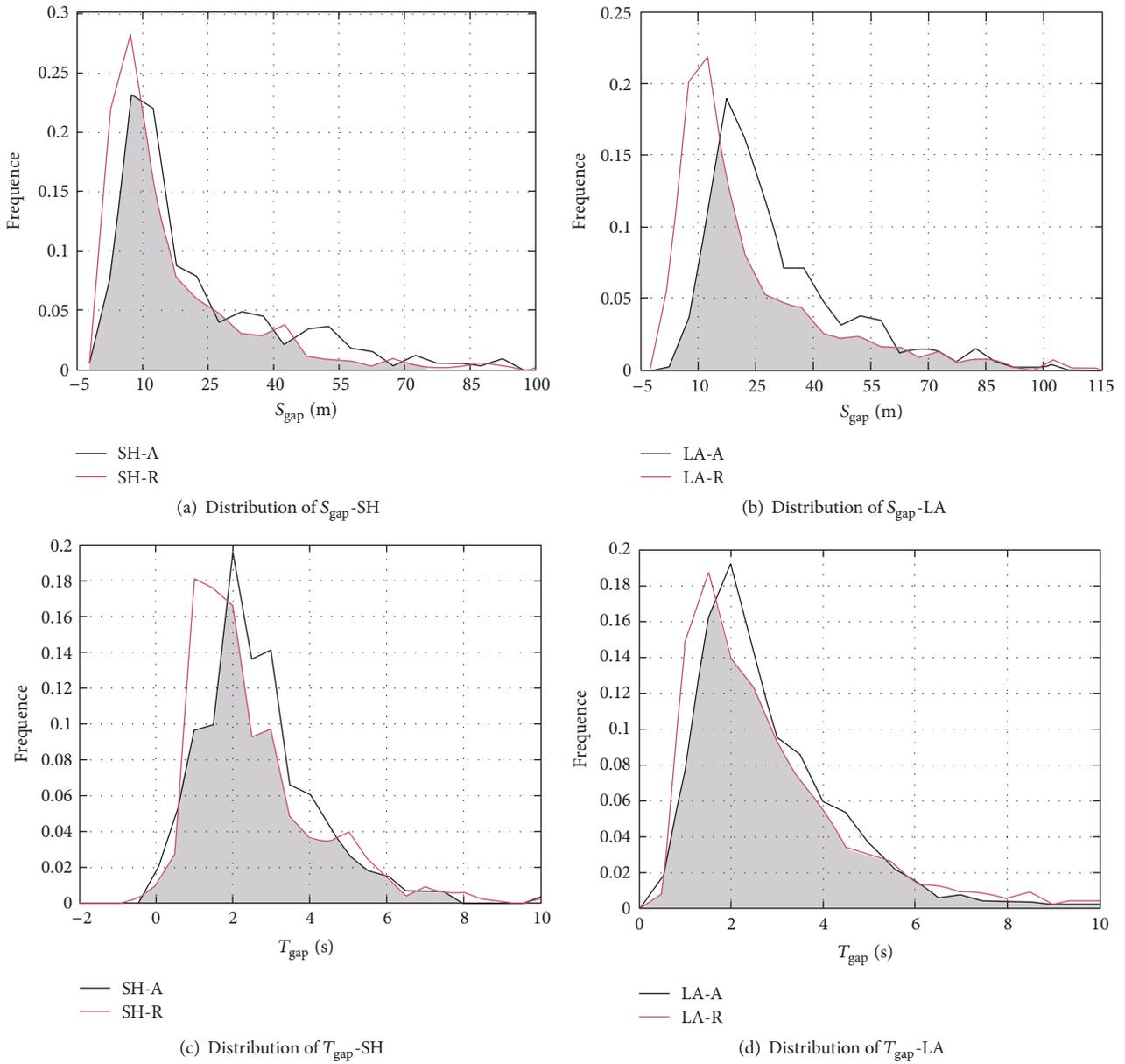


FIGURE 6: The distribution of S_{gap} and T_{gap} in SH and LA (note that SH-A and LA-A represent the accepted gaps in SH and LA, and SH-R and LA-R represent the rejected gaps.).

critical gap should also be treated as stochastic. This section focused on extending the concept and fitting the merging probability distribution function with respect to such critical gap when considering multi-rejected gaps.

5.1. Estimation of the Critical Gap. Survival analysis [31] is a branch of statistics for analyzing the expected duration of time until one or more events happen. It is widely used in medicine, biology, economics, and so on. Like stochastic capacity [32–34], the critical gap can also be estimated by using survival analysis. Previous [4] research shows that the time gap is more effective in predicting the merging behavior than the space gap. Therefore, in this study, T_{gap} is seen as the survival time (analogous to lifetime in lifetime data

analysis); the merging behavior corresponds to the failure event (analogous to death in lifetime data analysis); the rejected gap (not merge at a certain T_{gap}) is considered as censored data (analogous to the data that lifetime is longer than the duration of the experiment in lifetime data analysis) while the accepted gap (merge at a certain T_{gap}) is considered as uncensored data.

Product Limit Method (PLM), developed by Kaplan and Meier [35], is a non-parametric method of survival analysis that also includes the semi-parametric method and parametric method [31]. The non-parametric method is usually used to determine the survival probability under a certain survival time; the semi-parametric method is often chosen when analyzing the influence of variables on the survival

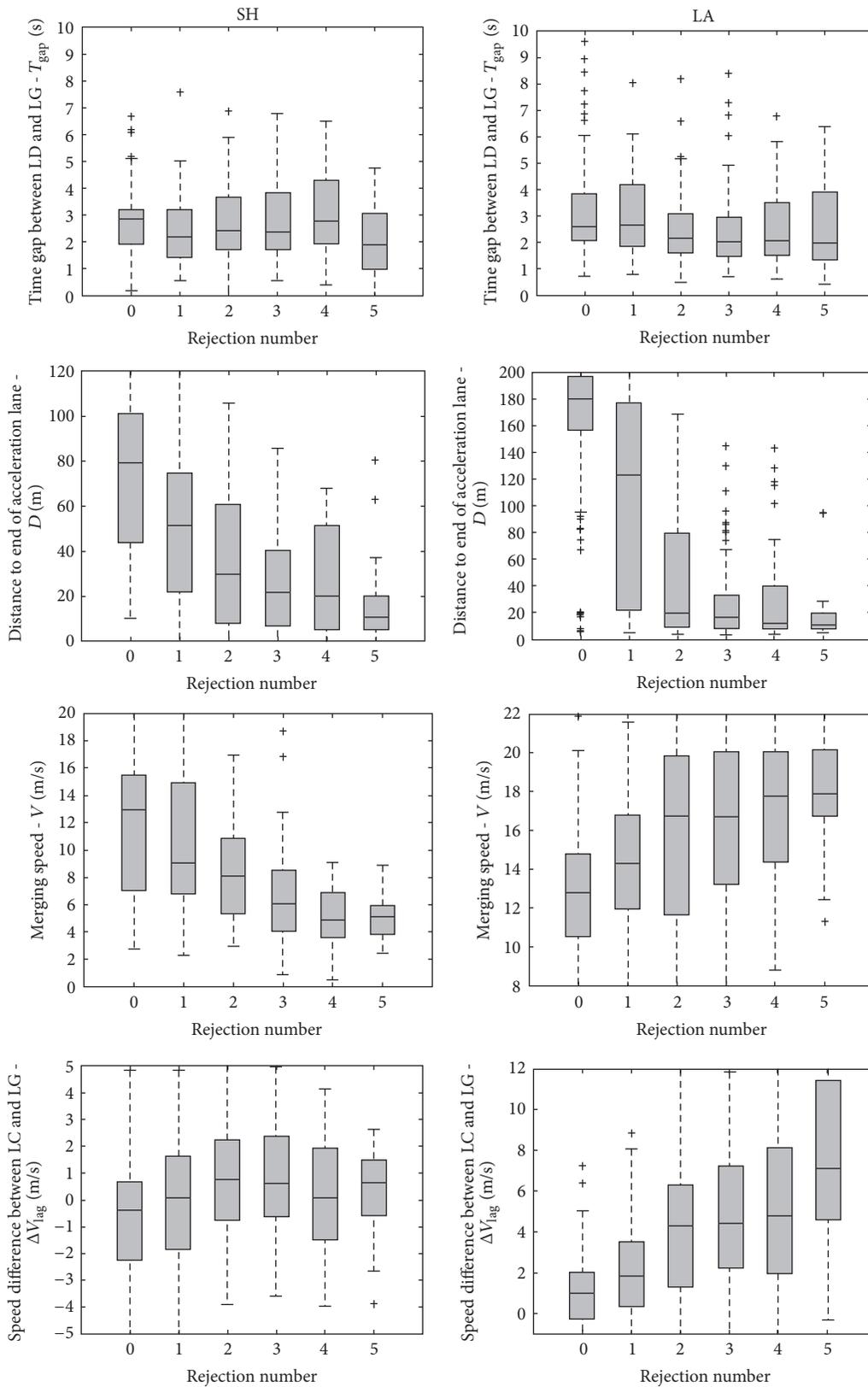


FIGURE 7: Merging features for various rejection numbers (the upper, middle, and lower lines of the box represent the 25th, 50th, and 75th percentiles of the data, resp.).

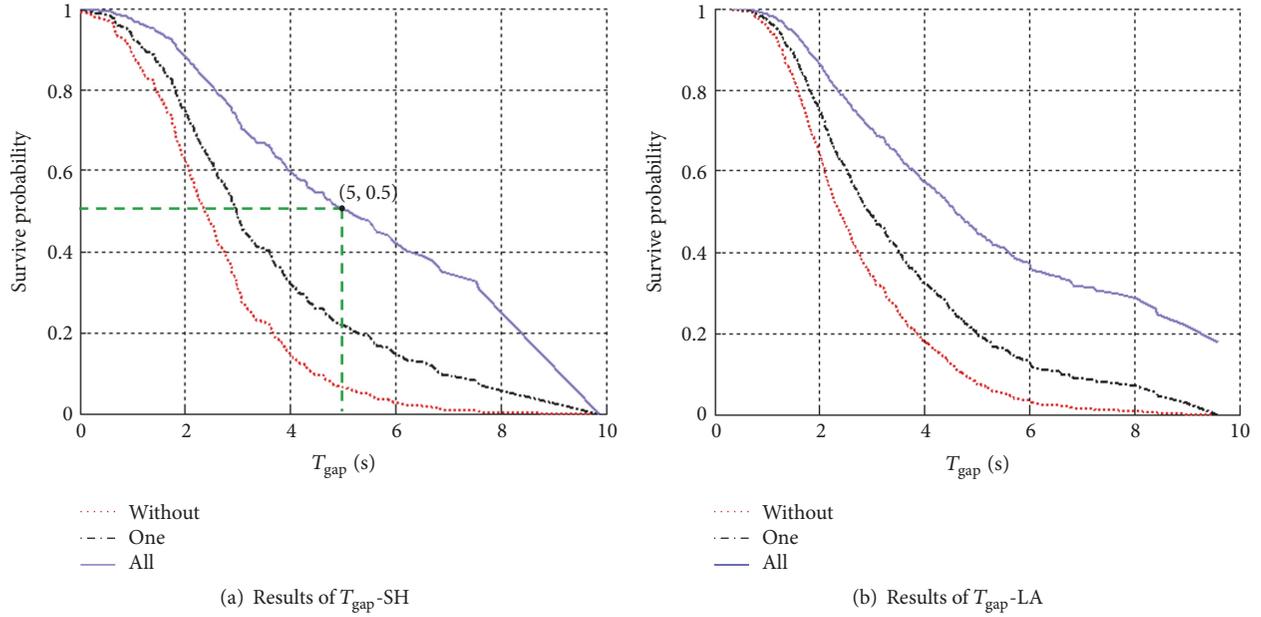


FIGURE 8: Results of PLM.

time; and the parametric method is based on the specific distribution of survival time and aims to build a parametric model of survival function. In this study, we firstly use the non-parametric PLM approach to get a distribution of survival probability. Then parametric PLM model is adopted after determining the distribution of survival probability.

PLM applies the multiplicative theorem of probability to calculate the survival probability. This calculation procedure is outlined below.

Step 1. Sort the survival time (i.e., T_{gap}) data in ascending order. Rank $i = 1, 2, 3, \dots, n$. If the value of censored data (i.e., rejected gaps) is equal to non-censored data (i.e., accepted gaps), the non-censored data is placed ahead.

Step 2. Calculate the survival probability of one rank starting from $i = 1$.

$$S_{(t)} = \prod_{i:t_i < t} \frac{n_i - d_i}{n_i}, \quad (2)$$

where t is survival time, that is, a certain T_{gap} ; S_t is the probability that survival time is longer than t , in other words, the probability that the vehicle did not merge at the time point t ; n_i is the number of vehicles with a longer survival time than t_i , that is, the number of vehicles that did not merge at a certain T_{gap} ; d_i is the number of death events at the time point t_i , that is, the number of vehicles that merged into the mainline at a certain T_{gap} .

Figure 8 shows the survival curve. It demonstrates that, under the same survival probability, the larger the rejection number considered, the larger the time gap. Meanwhile, as for the same time gap, the larger the rejection number considered, the higher the survival probability. For clarity and conciseness, *Without* represents the samples (gaps) with

only accepted gaps (uncensored data); *One* indicates the samples with uncensored data and censored data of one closest rejected gap before merging on time; *All* includes uncensored data and all censored data (rejected gaps). In Figure 8(a), for example, for the survival probability of 0.5, the time gap of *Without*, *One*, and *All* is 2.4 s, 2.9 s, and 5.0 s, respectively. For the time gap of 5.0 s, the survival probability of *Without*, *One* and *All* is 0.06, 0.22, and 0.50, respectively. These trends are the same for both the SH and LA sites. These analyses demonstrate that the rejecting behavior (reject a gap) has an impact on the accepted gaps, and comparing the results of considering just one rejected gap and multi-rejected gaps, the survival probability of the time gap is quite different. Therefore, multi-rejected gaps should be taken into account when estimating the critical gap rather than without thinking about it or just considering one largest rejected gap [36].

Furthermore, the slope of curve *All* in SH is bigger than that in LA, which shows a faster change of critical gap in SH. There are two likely reasons for this: (1) more vehicles in SH than in LA rejected at least one gap before merging (as demonstrated in Section 4.1); (2) there were a higher proportion of drivers in SH who rejected gaps that were bigger than the accepted gaps (as shown in Section 4.2).

5.2. Merging Probability Function with respect to Critical Gap. Let

$$F(t) = 1 - S(t), \quad (3)$$

where $F(t)$ is the probability of merging under a survival time t . In other words, it is the merging probability function with respect to critical gaps.

In order to use a parametric PLM, it is essential to determine the distribution of merging probability with respect to the critical gap.

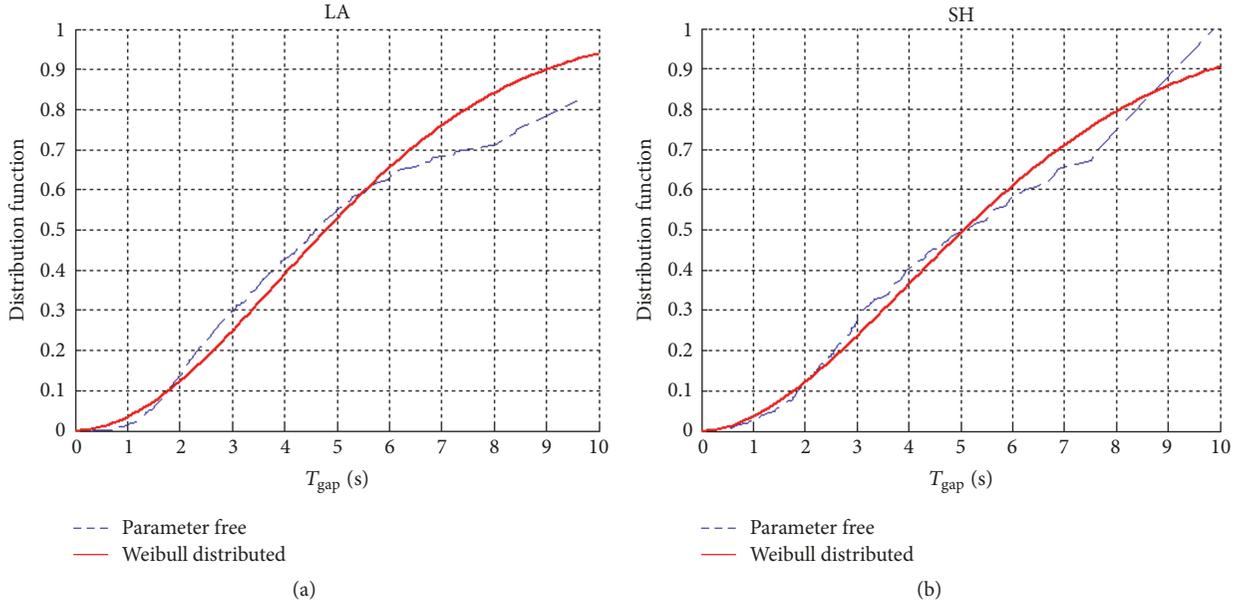
FIGURE 9: Weibull distribution function (the data used here are the samples of *All*).

TABLE 4: Results of parametric PLM.

Position	alpha	beta	$-2L$
LA	1.9	5.8	3452.4
SH	1.8	6.2	2025.1

Similar to stochastic capacity [32], based on the PLM result, various function types such as Normal, Lognormal, Weibull, exponential, and Log-logistic distribution curve are tested. We find that the Weibull distribution curve best performs the function of merging probability with respect to critical gap, whose function is expressed as follows:

$$F(x) = 1 - e^{-(x/\beta)^\alpha} \quad \text{for } x \geq 0, \quad (4)$$

where α is shape parameter; β is scale parameter.

When applying a parametric PLM, Maximum Likelihood Estimation (MLE) helps us to get an answer. The Maximum Likelihood Estimation function for a Weibull distribution is

$$L = \prod_{i=1}^n (a \cdot b^{-a} \cdot q_i^{\alpha-1} \cdot e^{-(q_i/b)^\alpha})^{\delta_i} \cdot (e^{-(q_i/b)^\alpha})^{1-\delta_i}. \quad (5)$$

Then by the means of genetic algorithm, we have a relative better result and it is listed in Table 4.

From Figure 9, we can see that the Weibull distribution well fits the merging probability distribution of the critical gap in both LA and SH. This result can be applied to microscopic traffic simulation, where the Weibull distribution could be more appropriate than the lognormal distribution [17]. Furthermore, the critical gap which is stochastic can help to explain free, forced, and cooperative lane changing [22, 23].

According to gap acceptance theory, we always accept the gap when it is larger than the traditional critical gap, and this leads to a free rather than a forced lane change. However, once we consider the stochastic nature of critical gap, the forced lane change could occur because the rejected gap is larger than the accepted gap. These forced lane changes might also lead to better simulation of early onset breakdown phenomena [2].

6. The Merging Behavior Prediction Model

In order to quantitatively measure how variables affect merging behavior, the utility-based logistic regression model is utilized to predict the merging behavior by considering the multi-rejected gaps.

6.1. Logistic Regression Model. In the logistic regression model, the merging utility is modeled as a function of explanatory variables affecting drivers' merging behavior. Maximum likelihood estimation is used to estimate parameters. The basic expression in the logistic regression model is shown in [37]

$$\text{logit}(y) = \ln\left(\frac{P}{1-P}\right) = \alpha + \sum_{i=1}^k \beta_i r_i, \quad (6)$$

where $\text{logit}(y)$ is the logarithm of the odds of experiencing an event (the linear relationship of the variables) and P is the probability of an event.

Because not every vehicle has a preceding vehicle, the explanatory variables T_{pre} and S_{pre} are excluded from the model, and the remaining variables (as shown in Table 1) are all taken into consideration. In our study, the confidence level is pre-set as 0.05. The final models (considering the multi-rejected gaps) for LA and SH are shown as follows:

$$LA: P = 1 - \frac{1}{\left[1 + \exp\left(3.566 + 0.159T_{lead} - 0.023D - 0.241\Delta V_{lag} + 0.12\Delta V_{lead}\right)\right]} \quad (7)$$

$$SH: P = 1 - \frac{1}{\left[1 + \exp\left(-0.481 + 0.51S_{gap} - 0.313T_{lead} + 0.121V - 0.028D + 0.345\Delta V_{lead} - 0.036S_{lag}\right)\right]}, \quad (8)$$

where P is the probability of merging into the mainline, with respect to the 1370 and 1964 samples (including both accepted and all the rejected gaps) in SH and LA, respectively.

The P value for parameters in (7) and (8) is lower than 0.05. And the Nagelkerke R -square is 0.337 for the logistic regression model in SH and 0.495 for that of LA. These two models well explain the merging behavior while the Nagelkerke R -square is higher than 0.15 [38].

In (7), T_{lead} , D , ΔV_{lag} , and ΔV_{lead} are the significant variables for describing the merging process in LA. The coefficients of T_{lead} and ΔV_{lead} are positive, which means that the larger T_{lead} or ΔV_{lead} is, the higher the merging probability becomes. The negative variables D and ΔV_{lag} is an indicator for the urgency of the merging. The smaller the remaining distance to the end of the merging point and bigger speed difference between merging behavior and putative follower, the higher the probability of merging into the mainline.

There are more variables included in (8) than in (7). This is mainly because of the more complex driving behavior and traffic conditions in SH than LA. There are three common variables in the SH and LA models: D , T_{lead} , and ΔV_{lead} . The sign of coefficients of D and ΔV_{lead} for SH is the same as that for LA; thus, the explanations are the same. The positive coefficient of the variable S_{gap} and V implies that there is a higher probability of a vehicle merging into the mainline when it meets a larger gap or owns a higher speed. Furthermore, the negative variables S_{lag} and T_{lead} could indicate a preference in SH for choosing a merging position near the leading vehicle on the target lane in time and near the following vehicle on the target lane in space. This means that the merging vehicles in SH prefer to overtake the following vehicle on the target lane and run after the leading vehicle on the target lane.

The prediction result for different rejection numbers is shown in Table 5, and is followed by a prediction accuracy comparison. A probability larger or equal to 0.5 is considered as a merging event, while a probability less than 0.5 is taken as a non-merging event.

In the following table, *One* represents the samples which include all the accepted samples and one closest rejected gap before merging; *Two* stands for the samples which contain all the accepted samples and two closest rejected gaps before merging, as well as *Three*, *Four*, and *All* (consisted of all the accepted and rejected samples).

The variables with * in Table 5 are significant in the logistic regression model. Table 5 demonstrates that variables have significant influences on merging behavior are different when considering various quantities of rejected gaps. As to SH, the *One* has four significant variables only covering the

merging vehicle and the lag vehicle at the target lane. As to the *All*, it owns six significant variables covering all the vehicles in a merging event, including time gap, space gap, and speed difference. Comparing the two models, the latter one concentrates on more factors and has a higher prediction accuracy. It implies that the behavior of rejecting a gap has impact on the subsequent decision, and the significant factor changes with the increasing of rejected gaps. Thus, multi-rejected gaps matter in the merging behavior. What is more, as to LA, *One* has five significant variables while *All* only has four. However, prediction accuracy of *All* is higher than *One*. It seems that considering multi-rejected gap contributes to the recognition of significant factors.

As shown in Table 6, the more the rejection number considered, the higher the prediction accuracy for both sites. This emphasizes the importance of multi-rejected gaps when predicting a merging behavior. What is more, relative to the prediction accuracy in SH, it is much higher in LA for the *One*, *Two*, *Three*, and *Four*. However, when considering all the rejected gaps, the prediction accuracy for both sites reaches 0.834, and the gap of prediction accuracy between two sites is filled. Since we have found in Section 4.1 that more vehicles in SH than in LA rejected at least one gap before merging, models of *One*, *Two*, *Three*, and *Four* neglect more rejected gaps in SH than in LA, which leads to a lower prediction accuracy in SH. When all the rejected gaps are considered (i.e., no rejected gaps are neglected), the prediction accuracy for both sites achieves improvement. To some extent, it further proves the necessity to consider multi-rejected gaps when analyzing and modeling merging behaviors.

6.2. Discussion. After introducing the multi-rejected gaps into the model, there is an improvement in accuracy for both sites. However, three related issues are worthy of discussion.

First, local impact factors such as road geometry influence the merging behavior. The previous work (Marczak et al. [7]) verified the conclusion. Actually, in our model, the conclusion is the same. In Table 5, the distance to the end of the acceleration lane D and the speed difference between the merging vehicle and the mainline vehicles ΔV_{lead} or ΔV_{lag} significantly influence the prediction of the merging behavior.

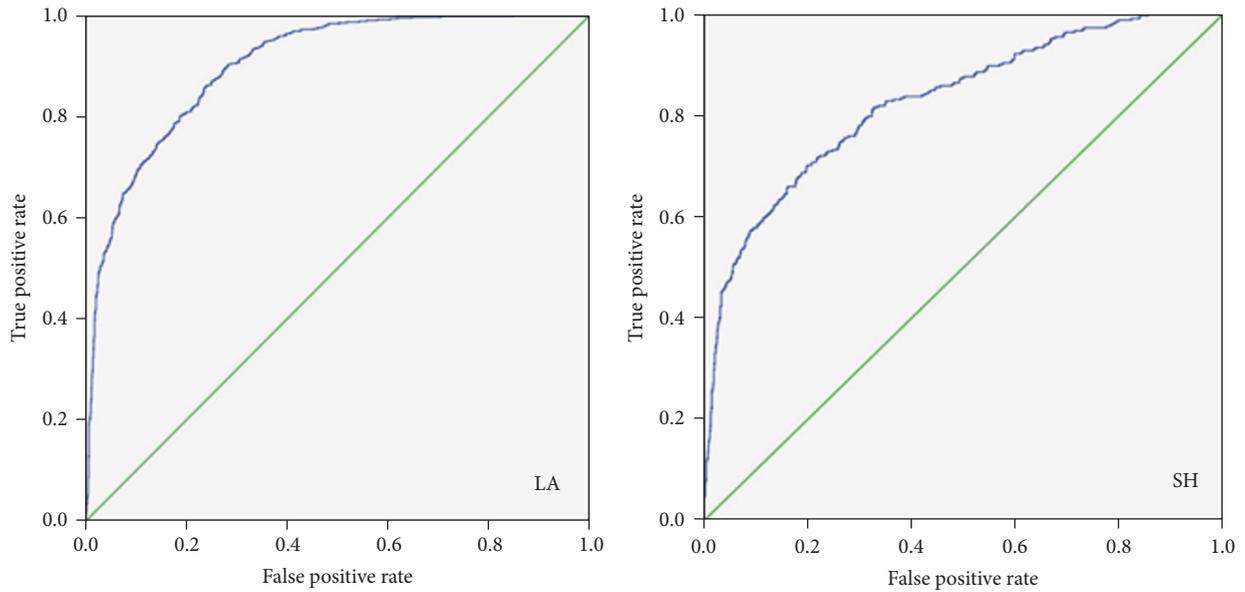
The prediction accuracy is still not very high compared with that found in other studies (86.8% in Hou et al. [10], and more than 98% in Marczak et al. [7]). Although NGSIM data set is also used in Hou et al. [10], the definition of “non-merging events” shows big difference in that study. Furthermore, there is a large overlap between accepted and rejected gaps (both the T_{gap} and S_{gap}) in our datasets; this makes it a more complex task to distinguish them and leads to

TABLE 5: Significant variables of logistic regression model.

Variables	SH					LA				
	One	Two	Three	Four	All	One	Two	Three	Four	All
D	*	*	*	*	*	*	*	*	*	*
V	*			*	*	*	*	*	*	
T_{gap}		*				*	*		*	
T_{lead}			*		*					*
T_{lag}		*						*		
S_{gap}			*		*	*	*	*	*	
S_{lead}							*	*	*	
S_{lag}	*	*	*		*	*	*	*	*	
ΔV_{lead}		*	*	*	*					*
ΔV_{lag}	*	*	*	*		*	*	*	*	*

TABLE 6: Prediction accuracy of logistic regression.

Rejection number	One	Two	Three	Four	All
LA	0.754	0.771	0.794	0.802	0.834
SH	0.709	0.721	0.735	0.758	0.834

FIGURE 10: ROC curves of two sites (true positive rate [sensitivity] represents the probability of predicting an accepted state as an accepted state, while false positive rate $[1 - \text{specificity}]$ is the probability of predicting a rejected state as an accepted state).

a relatively low accuracy. Meanwhile, Marczak et al. [7] used a different data set. We applied their method to our datasets, and the accuracy we obtained is only 75.4%.

The effectiveness of the prediction model can be verified by using the Receiver Operating Characteristic (ROC) curve [39]. As shown in Figure 10, under the same false positive rate, the true positive rate in LA is higher than that in SH. This indicates the more effective prediction of a merging event in LA and the lower probability of making the mistake of regarding accepted state as a rejected state. Furthermore, the point near the top-left corner is relatively more sensitive and

specific; thus, we can find the most appropriate threshold value to classify the merge and non-merge behavior. In this study, Youden's Index (sensitivity + specificity - 1) [40] is used to find the most appropriate threshold value. The largest Youden's Index is the threshold, we need to determine the most appropriate classification threshold, and they are 0.61 and 0.51 in LA and SH, respectively. Moreover, the area under the ROC curve (i.e., AUC) reflects the prediction effectiveness [41]. Generally, the AUC between 0.5 and 0.7 indicates a poor result; 0.7 to 0.9 indicate a moderate result; and the model performs well when the area is larger than 0.9. In this study,

the AUC for LA and SH is 0.902 and 0.827, respectively. From this, we can conclude that the model is effective.

7. Conclusions and Further Work

Based on the two trajectory data sets, one at Yan'an Expressway in Shanghai (China) and the other at Highway 101 in Los Angeles (USA), this study analyzed the merging behavior at the acceleration lane, especially the relationship between multi-rejected gaps and accepted gaps. From this analysis, we have drawn the following conclusions.

(1) The mean rejection number is 3.12 in SH and 2.25 in LA. By using t -test, both in SH and LA, the mean of accepted S_{gap} is significantly larger than the mean of rejected S_{gap} , while the mean of accepted T_{gap} is not significantly different to the rejected T_{gap} . Meanwhile, the result shows that the rejected gap can be larger than the accepted gap.

(2) The rejected gaps have impact on merging behaviors, and the merging preference is different between drivers in SH and LA. With the increase of rejection number, the merging condition becomes worse in SH while it is improved in LA. This result illustrates that drivers in SH are more risk-taking and self-focused while drivers in LA are more rational and altruistic.

(3) There is a significant difference in the critical gaps under different rejection numbers: the more the gaps are rejected, the larger the critical gap becomes. For example, for a merging probability of 0.5, with no rejected gaps considered, the critical gap is 2.4 s in LA and 2.8 s in SH; on the other hand, with all the rejected gaps considered, the critical gap is 5.0 s in SH and 4.5 s in LA. Meanwhile, survival analysis was undertaken to better understand the characteristics of the merging process, and the Weibull distribution function best fits the merging probability function of critical gap.

(4) Logistic regression models were developed to predict merging events. By taking into account multi-rejected gaps, the significant variables are more reasonable and efficient, and the prediction accuracy of merging behavior improves greatly. Comparing the model with only one closest rejected gap and all rejected gaps, there is a 10.61% and 17.63% improvement of the latter one in LA and SH, respectively. On the other hand, the local factors such as the road geometry influence the merging behavior. In the prediction model, the distance to the end of the acceleration lane and the speed difference between the merging vehicle and the mainline vehicles significantly influence the prediction of the merging behavior. Besides the randomness, the merging behavior may be dynamic because drivers may adjust the crucial gap after rejecting some ones, which is to be investigated in the future work.

This study focuses on the merging behavior in high-density flow, which has fundamental difference with the classical gap acceptance studies. In our study sites, speed of merging vehicles is larger and drivers are able to select a gap from several ones at the same time, which is a multinomial choice. For travel efficiency, the driver may prefer the gap downstream. For urgency, the driver may force in small gaps. Other factors outweigh the impact of gap. At a stop sign, the merging or crossing driver selects the oncoming gaps one by

one. Merging in low-density flow is in the same condition, because the speed of the merging vehicle is smaller than that of the mainline vehicles. Therefore, the classical gap acceptance theory for these scenarios is a binary choice. For whatever reasons such as efficiency and urgency, the best choice is to select the gap which is larger than the crucial one. Therefore, this study models a different merging behavior but does not imply that the gap acceptance behavior in other situations has the same properties.

Drivers in SH are more risk-taking and self-focused; this leads to more forced lane changings in SH and makes the modeling of merging behaviors for that site more challenging. This observation implies the necessity to adequately accommodate human factors in the microscopic modeling of driving behaviors (primarily, their lane-changing maneuvers and car-following behavior), as advocated in the recent literature (e.g., [5, 42, 43]). However, the explicit incorporation of risk perceptions in the modeling of the merging process was beyond the scope of this study and is a topic for future research.

Finally, drivers can experience several gaps during the merging process, and the gaps they finally accept are not necessarily the optimal ones. With the development of autonomous and connected vehicles, another area of future research is the optimization of the merging behavior decision model to ensure the operational efficiency of the merging area and the optimal and synchronized merging of all vehicles.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank the National Natural Science Foundation of China (U1764261 and 51422812) and the Science and Technology Commission, Shanghai Municipality (16DZ1203602), for supporting this research.

References

- [1] L. Leclercq, V. L. Knoop, F. Marczak, and S. P. Hoogendoorn, "Capacity drops at merges: New analytical investigations," in *Proceedings of the 2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, pp. 1129–1134, Qingdao, China, October 2014.
- [2] J. Sun, J. Zhang, and H. M. Zhang, "Investigation of early-onset breakdown phenomenon at urban expressway bottlenecks in Shanghai, China," *Transportmetrica B: Transport Dynamics*, vol. 2, no. 3, pp. 215–228, 2013.
- [3] M. J. Cassidy and R. L. Bertini, "Some traffic features at freeway bottlenecks," *Transportation Research Part B: Methodological*, vol. 33, no. 1, pp. 25–42, 1999.
- [4] J. Sun, L. Zhao, and H. Zhang, "Mechanism of early-onset breakdown at on-ramp bottlenecks on Shanghai, China, expressways," *Transportation Research Record*, vol. 2421, pp. 64–73, 2014.
- [5] Z. Zheng, "Recent developments and research needs in modeling lane changing," *Transportation Research Part B: Methodological*, vol. 60, pp. 16–32, 2014.

- [6] D. Chen, J. Laval, Z. Zheng, and S. Ahn, "A behavioral car-following model that captures traffic oscillations," *Transportation Research Part B: Methodological*, vol. 46, no. 6, pp. 744–761, 2012.
- [7] F. Marcjak, W. Daamen, and C. Buisson, "Merging behaviour: Empirical comparison between two sites and new theory development," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 530–546, 2013.
- [8] L. Zhao, J. Sun, and H. Zhang, "Observations and analysis of multistep-approach lane changes at expressway merge bottlenecks in shanghai, China," *Transportation Research Record*, no. 2395, pp. 73–82, 2013.
- [9] Z. Tian, R. Troutbeck, M. Kyte et al., "A further investigation on critical gap and follow-up time," in *Proceedings of the 4th International Symposium on Highway Capacity, Maui/Hawaii, Transportation Research Circular E-C018*, pp. 409–421, 2000.
- [10] Y. Hou, P. Edara, and C. Sun, "A genetic fuzzy system for modeling mandatory lane changing," in *Proceedings of the 2012 15th International IEEE Conference on Intelligent Transportation Systems, ITSC 2012*, pp. 1044–1048, Anchorage, Alaska, USA, September 2012.
- [11] Q. Meng and J. Weng, "Classification and regression tree approach for predicting drivers' merging behavior in short-term work zone merging areas," *Journal of Transportation Engineering*, vol. 138, no. 8, pp. 1062–1070, 2012.
- [12] W. Daamen, M. Loot, and S. P. Hoogendoorn, "Empirical analysis of merging behavior at freeway on-ramp," *Transportation Research Record*, no. 2188, pp. 108–118, 2010.
- [13] G. A. Davis and T. Swenson, "Field study of gap acceptance by left-turning drivers," *Transportation Research Record*, vol. 1899, pp. 71–75, 2004.
- [14] P. Lertworawanich and L. Elefteriadou, "A methodology for estimating capacity at ramp weaves based on gap acceptance and linear optimization," *Transportation Research Part B: Methodological*, vol. 37, no. 5, pp. 459–483, 2003.
- [15] M. A. Pollatschek, A. Polus, and M. Livneh, "A decision model for gap acceptance and capacity at intersections," *Transportation Research Part B: Methodological*, vol. 36, no. 7, pp. 649–663, 2002.
- [16] V. Goswami and H. Ghulam, "Gap acceptance behavior in mandatory lane changes under congested and uncongested traffic on a multilane freeway," in *Proceedings of the Transportation Research Board 86th Annual Meeting*.
- [17] T. Toledo, H. N. Koutsopoulos, and M. E. Ben-Akiva, "Modeling integrated lane-changing behavior," *Transportation Research Record*, no. 1857, pp. 30–38, 2003.
- [18] A. Halati, H. Lieu, and S. Walker, "CORSIM- corridor traffic simulation model," in *Proceedings of the 1997 Conference on Traffic Congestion and Traffic Safety in the 21st Century*, pp. 570–576, June 1997.
- [19] V. PTV, 5.30 User Manual. PTV Planung Transport Verkehr AG, Stumpfstraße 1, D-76131 Karlsruhe, 2010.
- [20] Caliper, "TransModeler Traffic Simulation Software – Version 2.5 User's Guide," 2009.
- [21] Z. Zheng, S. Ahn, D. Chen, and J. Laval, "The effects of lane-changing on the immediate follower: anticipation, relaxation, and change in driver characteristics," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 367–379, 2013.
- [22] P. Hidas, "Modelling vehicle interactions in microscopic simulation of merging and weaving," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 1, pp. 37–62, 2005.
- [23] P. Hidas, "Modelling individual behaviour in microsimulation models," in *Proceedings of the 28th Australasian Transport Research Forum, ATRF 2005*, Sydney, Australia, August 2005.
- [24] F. C. Fang and L. Elefteriadou, "Some guidelines for selecting microsimulation models for interchange traffic operational analysis," *Journal of Transportation Engineering*, vol. 131, no. 7, pp. 535–543, 2005.
- [25] P. G. Gipps, "A model for the structure of lane-changing decisions," *Transportation Research Part B: Methodological*, vol. 20, no. 5, pp. 403–414, 1986.
- [26] D. Sun and L. Elefteriadou, "Information categorization based on driver behavior for urban lane-changing maneuvers," *Transportation Research Record*, no. 2249, pp. 86–94, 2011.
- [27] J. Sun, J. Ouyang, and J. Yang, "Modeling and analysis of merging behavior at expressway on-ramp bottlenecks," *Transportation Research Record*, vol. 2421, pp. 74–81, 2014.
- [28] Y. Hou, P. Edara, and C. Sun, "Modeling mandatory lane changing using bayes classifier and decision trees," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 647–655, 2014.
- [29] Z. Zheng, S. Ahn, D. Chen, and J. Laval, "Freeway traffic oscillations: Microscopic analysis of formations and propagations using Wavelet Transform," *Transportation Research Part B: Methodological*, vol. 45, no. 9, pp. 1378–1388, 2011.
- [30] J. Migletz, J. L. Graham, I. B. Anderson, D. W. Harwood, and K. M. Bauer, "Work Zone Speed Limit Procedure," *Transportation Research Record*, no. 1657, pp. 24–30, 1999.
- [31] D. G. Kleinbaum and M. Klein, *Survival Analysis*, Springer, New York NY, USA, 1996.
- [32] W. Brilon, J. Geistefeldt, and M. Regler, "Reliability of freeway traffic flow: a stochastic concept of capacity," in *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, pp. 125–143, College Park, Md, USA, 2005.
- [33] W. Brilon, J. Geistefeldt, and H. Zurlinden, "Implementing the concept of reliability for highway capacity analysis," *Transportation Research Record*, vol. 2027, no. 1, pp. 1–8, 2007.
- [34] V. Modi, A. Kondyli, S. S. Washburn, and D. S. McLeod, "Freeway capacity estimation method for planning applications," *Journal of Transportation Engineering*, vol. 140, no. 9, Article ID 05014004, 2014.
- [35] E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *Journal of the American Statistical Association*, vol. 53, pp. 457–481, 1958.
- [36] N. Wu, "Equilibrium of probabilities for estimating distribution function of critical gaps at unsignalized intersections," *Transportation Research Record*, no. 2286, pp. 49–55, 2012.
- [37] J. M. Hilbe, *Logistic Regression Models*, Chapman & Hall/CRC Press, Boca Raton, Fla, USA, 2009.
- [38] S. Menard, *Applied Logistic Regression Analysis*, SAGE Publications, 2002.
- [39] B. J. McNeil and J. A. Hanley, "Statistical approaches to the analysis of receiver operating characteristic (ROC) curves," *Medical Decision Making*, vol. 4, no. 2, pp. 137–150, 1984.
- [40] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.
- [41] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
- [42] M. Saifuzzaman and Z. Zheng, "Incorporating human-factors in car-following models: a review of recent developments

and research needs,” *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 379–403, 2014.

- [43] M. Saifuzzaman, Z. Zheng, M. Mazharul Haque, and S. Washington, “Revisiting the task-capability interface model for incorporating human factors into car-following models,” *Transportation Research Part B: Methodological*, vol. 82, pp. 1–19, 2015.



Hindawi

Submit your manuscripts at
www.hindawi.com

