Discretionary lane-changing behavior:

empirical validation for one realistic rule-based model

Cheng-JieJin[1,2,3], Victor L. Knoop [3], Dawei Li[1,2], Ling-Yu Meng[1,2], Hao Wang[1,2]

[1] Jiangsu Key Laboratory of Urban ITS, Southeast University of China, Nanjing, Jiangsu, 210096, People's Republic of China

[2] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University of China, Nanjing, Jiangsu, 210096, People's Republic of China

[3]Delft University of Technology, Transport & Planning, Stevinweg 1, 2628 CN Delft, The Netherlands

Abstract:

In this paper we discuss the mechanisms for discretionary lane-changing behavior in traffic flow. NGSIM video data is used to check the validity of different lane-changing rules, and 373 lane changes at 4 locations in US-101 highway are analyzed. We find the classical lane-changing rules of rule-based model, cannot explain many cases in the empirical dataset. Therefore, we propose one new decision rule, comparing the position after a time horizon of several seconds without a lane-change. This rule can be described as "to have a further position within 9 seconds". The tests on NGSIM data show that this rule can explain most (76%) of the lane-changing cases. Besides, some data when lane changes do not occur is also studied. We find most (81%) of non-lane-changing vehicles do not fulfill the new rule. Thus it can be considered as one sufficient and necessary condition for discretionary lane-changing.

Keywords: discretionary lane change, lane-changing model, rule-based model, NGSIM data, non-lane-changing vehicles

1. Introduction

For the modeling of traffic flow, usually there are two important topics for study: the car-following in the single-lane traffic, and the lane-changing in the multi-lane traffic. Lane-changing behavior has significant effects on the traffic operations, and often become the source of traffic jams. Usually, there are two different types of lane changes: the discretionary lane changes (DLCs) when drivers want to change lanes due to traffic conditions, and the mandatory lane changes (MLCs) when drivers need to change lanes in order to reach their desired destinations, including merging and diverging behavior. In this paper we focus on the former ones.

There are many different attempts to modeling various lane changes (Moridpour et al., 2013; Rahman et al., 2013; Zheng, 2014), including the macroscopic models (Sheu and Ritchie, 2001; Laval and Leclercq, 2008; Jin, 2010; Zheng et al., 2013) and microscopic models. For the microscopic models ones, various approaches are used, including: (1) the rule-based models (Gipps, 1986; Yang and Koutsopoulos, 1996; Sun and Elefteriadou, 2010; Sun and Kondyli, 2010; Sheu, 2013), in which the lane-changing reasons are evaluated first. If these reasons warrant a lane change, a gap acceptance model will be used to determine whether the gaps should be accepted. (2) the discrete-choice-based models (Ahmed et al., 1996; Toledo et al., 2007;

Choudhury and Ben-Akiva, 2013), in which all the steps are based on logit or probit models. (3) the artificial intelligence models, including fuzzy-logic-based models (Wu et al., 2003) and artificial neural network models (Hunt and Lyons, 1994). (4) the incentive-based models, including MOBIL (Kesting et al., 2007) and LMRS (Schakel et al., 2012), etc. Here drivers try to maximize their benefits, and the decision is based on the comparison between the "advantage" value and the threshold value. Among them, the rule-based models are very easy to understand and use, which will be the main topic of this paper.

In the previous rule-based models, usually emphasis is put on the instant status of the adjacent vehicles when discretionary lane changes occur, including their positions, gaps, velocities and accelerations, etc. Many different equations and parameters are used, and some complex ideas such as gap acceptance or time-headway distributions are involved. However, the validities of these models are not clear, due to the lack of enough empirical data of lane-changings. Many simulations have been run, but we do not know whether they are useful to the real life.

There are some studies for the calibration and validation of other lane-changing models, in which empirical traffic data is used (Yang and Koutsopoulos, 1996; Leclercq et al., 2007; Thiemann et al, 2008; Yeo and Skabardonis, 2008; Yeo et al., 2008; Moridpour et al., 2010; Schakel et al., 2012; Knoop and Buisson, 2015; Park et al., 2015; Lee et al., 2016), but they also have shortcomings. For example, some use the vehicle trajectories in one lane, but do not consider what happened between the vehicles in different lanes; some present many useful microscopic results, but lack an integrating framework for lane changes. Particularly, the methods introduced in these papers cannot be directly used for the rule-based models.

Therefore, in this paper we propose a new idea about the mechanisms of DLCs with rule-based models, and we think the new rules may improve the usefulness of rule-based models. Section 2 elaborates on the empirical NGSIM dataset we used. Here we choose the cellular automaton (CA) models as typical examples, and check the validity of some classical lane-changing rules in CA models (Wagner et al., 1996; Chowdhury et al., 1997; Nagel et al., 1998) with NGSIM dataset. Section 4 presents the new rules, taking the future conditions into account. Then the lane changes in the dataset are classified into several groups. In Section 5, different groups are used to test the lane-changing rules respectively. We find the new rules are more realistic, and they perform much better than the classical ones. In Section 6, we try to study some data when the lane changes do not occur, which is seldom done in many previous studies. Finally, the conclusion is given in Section 7.

2. The empirical data of lane changes

In order to study the details of DLC, empirical data with enough detail of lane changes are needed. Video data is better suited than loop detector data, since we are able to track lanes, and hence lane changes for individual vehicles. For this purpose, NGSIM data (FHWA, 2008; Delpiano et al., 2015) is a good choice. The high-quality video data gives us many details of the highway traffic flow, and can be freely downloaded from the Internet. All the related factors of lane changes, including the velocities of vehicles and the gaps between different vehicles can be obtained and used. The data of US-101 highway and I-80 highway are both possible for this study, but here we only use US-101 data, since the HOV lane in I-80 highway makes the lane-changing behaviors much more complex. Besides, many locations of I-80 highway are near the ramps, which implies the possible existence of many MLCs. We cannot easily judge whether the lane changes observed at these locations belong to DLC or not, thus we choose to abandon them.

In the data of US-101 highway, there are 8 cameras which can be used, as shown in Fig.1. There are 5 lanes on this highway, and the leftmost lane is marked as Lane 1, while the rightmost one is Lane 5. Here we only use the data of Camera 1, 2, 7, 8, since the locations of Camera 3, 4, 5, 6 are close to the on- and off-ramps. At the 4 locations, the lane changes from 7:50 to 8:20 on Jun.15, 2005 are observed and analyzed. Since there are some errors in the existing trajectory data of NGSIM video (Wang et al., 2014), we also use one software named Tracker (https://physlets.org/tracker/) to extract the lane-changing data. In the process, we use manual checks to avoid the errors, ensuring a higher robustness than that with a fully automated process.

It should be noted that some lane changes are excluded from our study due to the following reasons:

(1) The ones when some important factors cannot be recorded. For example, in this video data all the vehicles run from right to left. If the lane changes just occur at the left edge of the video, the front gaps on the previous lane and the target lane may be lost.

(2) The consecutive lane changes, i.e., two lane changes of one vehicle occur within very short time. For this situation, the vehicle may become very aggressive and the mechanism may be different. In this case, all lane changes, including the first, are ignored

(3) Two adjacent vehicles may change lanes at the same time. For this complex situation we only record the lane-changing data of the most downstream vehicle.

Fig.1. The study area of US-101 highway, provided by FHWA reports (http://www.ngsim.fhwa.dot.gov).

In some recent studies using NGSIM data (Park et al., 2015; Lee et al., 2016), all the lane changes in which the vehicle moves to the right lane are excluded, and they are all simply considered as MLCs. But in this paper they are all included, since we find many of them could be DLCs, e.g., they can be explained by the simple rules of DLC. We think the direction of lane changes may be not so important, which will be discussed later in Section 6.

Therefore, there are N_L=373 lane changes which can fulfill our criterion, and they become the basic dataset of this paper.

3. The classical lane-changing rules and their validity

3.1. The classical rules

In the previous studies of DLC behaviors with rule-based models, the instant status of vehicles when lane-changing occurs are important. As shown in Fig.2, the related factors are:

vehicle A: the lane-changing vehicle;
vehicle B: the front vehicle in the previous lane;
vehicle C: the front vehicle in the target lane;
vehicle D: the back vehicle in the target lane;
vehicle E: the back vehicle in the previous lane;
V0: velocity of vehicle A;
V1: velocity of vehicle B;
V2: velocity of vehicle C;
V3: velocity of vehicle D;
V4: velocity of vehicle E;
G1: gap between vehicle A and vehicle B;
G2: gap between vehicle A and vehicle C;
G3: gap between vehicle A and vehicle D;
G4: gap between vehicle A and vehicle E;

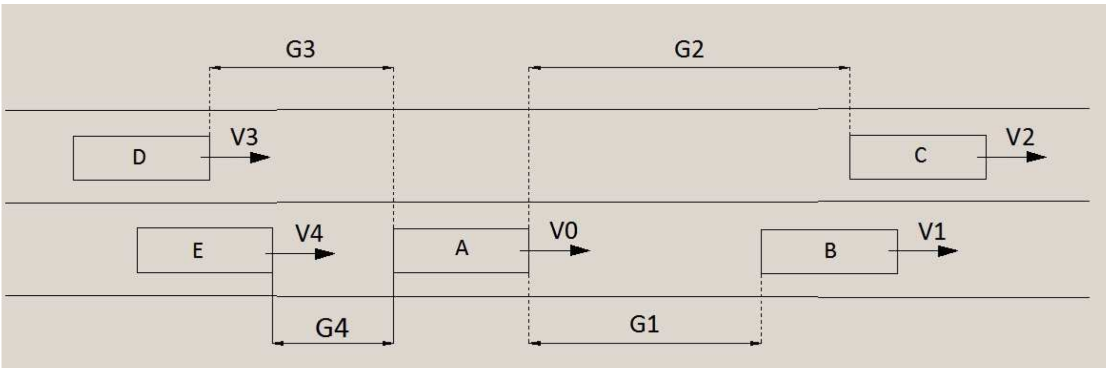


Fig.2. The schematic illustration of typical DLC behaviors.

For the sake of simplicity, we take CA models as examples, but the overall reasoning would hold for other rule-based models as well. In CA models the time and space are both discretized, and the time step is usually T=1s (Wolfram, 1983; Nagel and Schreckenberg, 1992). In the classical studies of lane-changing behaviors in the 1990s (Wagner et al., 1996; Chowdhury et al., 1997; Nagel et al., 1998), usually lane changing occurs if all of the following conditions are met[1]:

Condition 1: $V_0 > G_1$ (the movement of vehicle A in the next time step will be hindered);

Condition 2: $G_2 > G_1$ (the target lane has more room than the current lane);

Condition 3: $G_3 > V_3$ (the lane-changing behavior will not affect the movement of back vehicle);

1 In these equations, we scale the speeds to gaps by assuming T=1. This can be done without loss of generality by choosing the appropriate unit for time. Besides, the unit of all the velocities presented in this paper is m/s.

And there are various complex forms in the following studies (Knospe et al., 1999, 2002;Jia et al., 2005; Li et al., 2006; Kukida et al., 2011; Hu et al., 2012). For example:

(1) $V_0 > G_1 + V_1$ or $V_3 < G_3 + V_0$ (vehicle A considers the potential velocity of vehicle B, or vehicle D considers that of vehicle A);

(2) $G_2 > G_1 * t$ (t>1,which means there should be enough room to stimulate vehicle A to change lanes, and t can be adjusted);

(3) The probability of lane-changing (P) is introduced. Usually there is 0<P<1, e.g., P=0.5 or 0.2 for some cases, in order to eliminate the phenomenon of "ping-pong lane changes".

But the basic concept remains the same. The core concepts of these rules is: the decision of DLC is based on the instant status of surrounding vehicles. The benefit of lane-changing can be obtained in the next time step, i.e., immediately after the execution of lane changes.

3.2. The validity in empirical data

Since all the needed factors can be obtained in the NGSIM data, it is easy to use these data to check the validity of these classical rules. However, we find the results for the classical lane-changing rules are not good, which can be clearly seen in the following distributions.

Firstly, we use the calculations of $G_1 - V_0$ to check Condition 1. It should be noted that in all the distributions, the values are the proportions between the two scales of the X-axis. For example, In Fig.3 the interval is 3m, and the data of "12" represents the result when $9\text{m} \le G_1 - V_0 < 12m$. It is clear that many lane changes do not fulfill Condition 1, and the Effective Proportion (**EP**s for short) is only about 45% for $G_1 - V_0 < 0$. Especially, the peak in Fig.3 corresponds to $0 \le G_1 - V_0 < 3m$, which implies most vehicles decide to change lanes when the gap is a little larger than expected. Thus there is no need to study the validity of $V_0 > G_1 + V_1$, which is more difficult to fulfill.

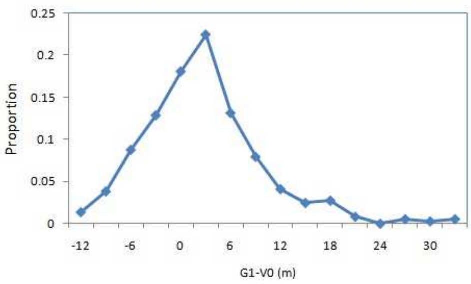


Fig.3. The distribution of $G_1 - V_0$ in the data of lane changes, which corresponds to Condition 1.

Then we consider the Condition 2, and the results[2] of the calculations of $G_1 - G_2$ are shown in Fig.4. We find the results are similar to Fig.3, since many lane changes do not fulfill the Condition 2, and the EP is even lower: about 38% for $G_1 - G_2 < 0$. The peak in Fig.4 corresponds to $4\text{m} \le G_1 - G_2 < 8m$, which means most vehicles decide to change lanes when the gap in the target lane is a little smaller. This phenomenon is completely different from what we expected before, and then, there is no need to study the validity of $G_2 > G_1 * t$ (t>1).

2 There are some special lane-changing cases in which G2<0, and we do not consider them in Fig.4. The discussion of these cases can be found in Section 5.

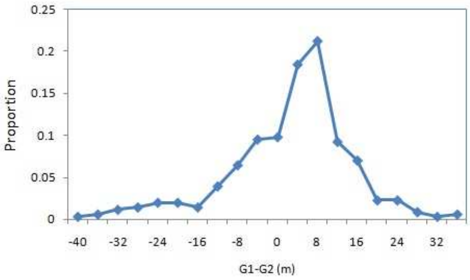


Fig.4. The distribution of $\mathrm{G_1 - G_2}$ in the data of lane changes, which corresponds to Condition 2.

Finally we consider the results of Condition 3, and the calculations of $\mathrm{V_3 - G_3}$ are shown in Fig.5. The EP seems higher than that in other two conditions: it is about 60% for $\mathrm{V_3 - G_3} < 0$. We can say more than half drivers consider the situations of the others behind when they change lanes. Nevertheless, this condition is not met by many other drivers, e.g., the peak in Fig.5is found at $0 \leq \mathrm{V_3 - G_3} < 3m$.

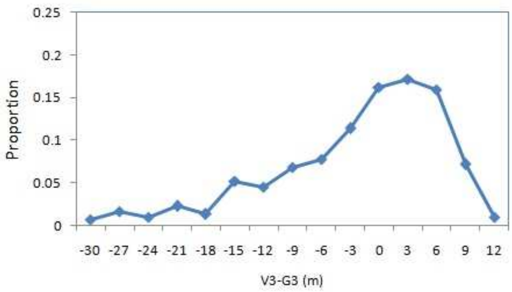


Fig.5. The distribution of $\mathrm{V_3 - G_3}$ in the data of lane changes, which corresponds to Condition 3.

If we consider the combined result of three classical rules, the EP becomes even lower: only. This further shows the irrationality of these rules. In a word, we find many vehicles choose to:

(1) Change lanes when the movement in the next time step is not really hindered. Sometimes the current gap ($\mathrm{G_1}$) is large, but the vehicles do not want to wait.

(2) Change lanes when the gap in the adjacent lane is not larger than the current gap. These changes are hard to reason from a traditional point of view, but this happens in about 40% of the cases.

(3) Change lanes when the back vehicle in the adjacent lane will be affected. Sometimes it will be seriously affected, since in some cases $\mathrm{G_3}$ is close to 0, or even negative. In these situations, the vehicles are very radical and just want to overpass the others.

These phenomena clearly shows that the benefit of lane changes cannot be immediately obtained. We need some other realistic rules which describe the lane changes better, and these rules are introduced in the next section.

4. The new lane-changing rules

In the model, we propose the motivation for DLC should be "to move faster in the future", i.e., can have a further position after some time horizon. This model can be presented by one simple equation as follows:

$$X_{c,T} > X_{n,T} \quad (1)$$

Here $X_{c,T}$ is the vehicle position after time T when it changes lanes, and $X_{n,T}$ is the position after time T when it does not change lanes. T is a time horizon which will be investigated later, by varying the value of T and checking which part of the lane changes is explained by the model. This rule is very easy to understand. It implies the decision for lane-changing depends on the estimation results of the traffic in the two lanes, rather than the current status of the adjacent vehicles.

If we set T=0s, this equation will degenerate to one of the classical rules (Condition 2), which has been proved to be invalid. Thus the following task is to find one realistic method to calculate the proper value of T, especially from the perspective of drivers. One possible way is to observe the front vehicles on the two lanes, and calculate their future positions. Then we change (1) to (2):

$$G_{1,T} < G_{2,T} \quad (2)$$

Here $G_{1,T}$ and $G_{2,T}$ are the estimated front gaps on the two lanes after time T. This new equation is different from Rule 2, since $G_{1,T}$ and $G_{2,T}$ are determined not only by the current gaps (G_1 and G_2), but also by future velocities. Since the drivers only know the current velocities of other vehicles, here we use their current velocities (V_1 and V_2) to do the calculation. Thus we change (2) to (3):

$$G_1 + V_1 * T < G_2 + V_2 * T \quad (3)$$

So the critical value of T for this lane change is:

$$T_a = \frac{G_2 - G_1}{V_1 - V_2} \quad (4)$$

The Equation (4) is simple, and easy to be checked by the lane-changing data. In the following study we only use Equation (4). Note that in some other rule-based models, the accelerations of vehicles are also considered. But we think it is not easy to estimate the future accelerations of neighboring vehicles in the empirical data, especially when the time horizon is large. Even we suppose it is constant during the following time, we do not know when this process ends. Thus in our model, we do not consider the effect of accelerations.

Then we can divide all the lane-changing data into four groups:

Group A: $G_2 > G_1$ and $V_1 < V_2$. This is the "best" condition for lane-changing, and there is no need to analyze the data which belong to this group.

Group B: $G_2 > G_1$ and $V_1 \geq V_2$. This is one special situation, in which $G_{2,T}$ will decrease and may be smaller than $G_{1,T}$ in the future.

Group C: $G_2 \leq G_1$ and $V_1 < V_2$. This is very important for our study, since the future movement of the front vehicles need to be evaluated by the drivers of lane-changing vehicles.

Group D: $G_2 \leq G_1$ and $V_1 \geq V_2$. This is the "worst" condition for possible lane-changing, and the lane changes in this group are difficult to understand, since the benefits from these lane changes are not clear. But we also try to explain them later.

We find the numbers of lane changes which belong to Group A, B, C, D are N_A=112, N_B=32, N_c=161 and N_D=68 in our dataset. For Group B, C, D, the overall proportion is about 70%, which needs to be further discussed. Here we introduce the critical parameter (T_c) for Group B and C. We suppose there could be one T_c for all the drivers, and when equation (5) is met, they choose to change lanes:

$$\begin{cases} \text{Group B: } T_a \geq T_c \\ \text{Group C: } T_a < T_c \end{cases} \quad (5)$$

Here T_c also means the anticipation time of drivers. After T_c, the drivers can benefit from the lane changes.

5. The validity of new lane-changing rules

In this section, the quality of the newly proposed rules is tested, by comparing the decisions predicted by the rules with the actual decisions observed in the video data. Firstly, the results for Group B and C are shown in Fig.6. Here the values on X-axis mean the maximum ones, e.g., the value of "$T_a = 3s$" corresponds to the proportion when $T_a < 3s$. The basic tendencies of two cumulative curves in Fig.6are similar, but the growth of the proportion in Group C is much faster than that in Group B. For example, in Group B about 91% of lane changes have $T_a < 9s$, but in Group C the result is only about 66%.

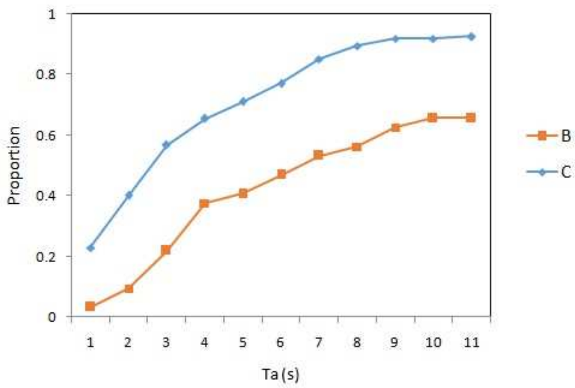


Fig.6. The cumulative curve of T_a values in Group B and C of lane-changing vehicles.

Then we consider the combined EPs of two groups. Note that the rules for Group B and C are different($T_a \geq T_c$and $T_a < T_c$), thus the results should be:

$$EP_{B+C} = \frac{P_c * N_c + (1 - P_B) * N_B}{N_B + N_c} \quad (6)$$

Where P_B and P_c are the proportions shown in Fig.6, N_B=32 and N_C=112. Then the cumulative curve of EPs is shown in Fig.7. Since $N_C > N_B$, the combined results are mainly determined by Group C. Here we think $T_c = 9s$ can be considered as one critical value in the empirical data, and the corresponding EP_{B+C} is about 83%. On one hand, when $T_c < 9s$ the EPs monotonically increase, and when $T_c > 9s$, they seldom increase and keep nearly constant. On the other hand, in both Group B and C, the numbers of the cases in which $T_c > 9s$ are small. We call the N_{BC1}=160 cases in which Equation (5) is fulfilled as "Group BC1", and the other N_{BC2}=33 cases as "Group BC2".

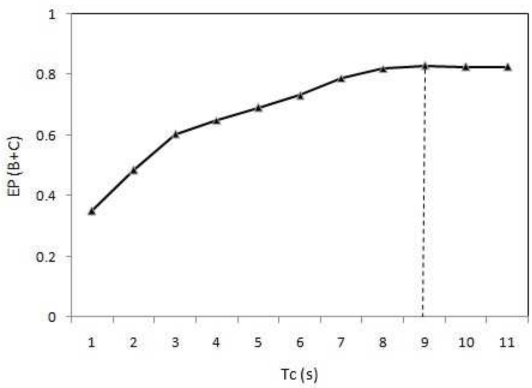


Fig.7. The EPs of new rules in Group B+C of lane-changing vehicles.

Then the data of Group D and Group BC2 need to be further investigated. We call these 101 cases as Group X, and it can be divided into two groups:

X1: Give way to others. Here we have N_{X1}=13. There are two different situations:

(1) To avoid the influence of large trucks. There are 4 cases which have relationship with the nearby large trucks. They can be clearly identified in the video.

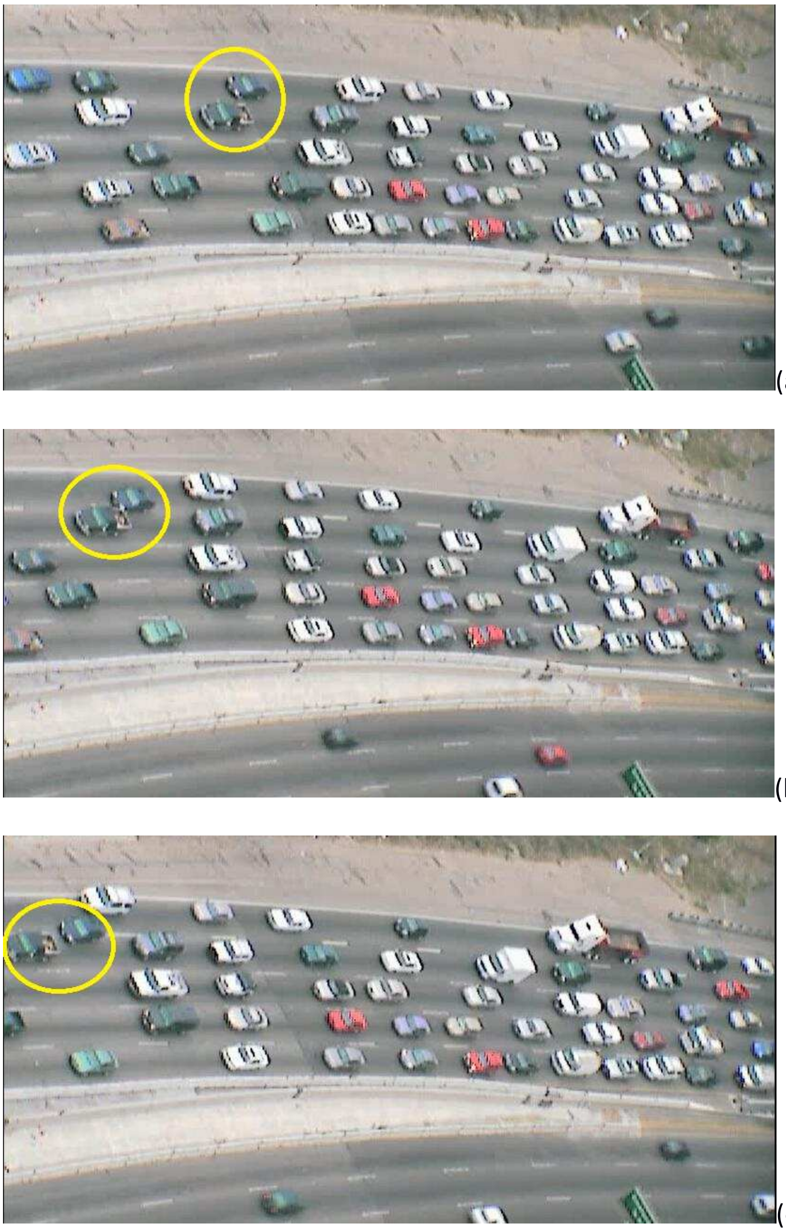
(2) To give way to the following vehicle. There are 9 cases which have relationship with the following vehicle. Here it is difficult and not necessary to build one new model with only 9 cases, thus we use one simple way to check. In the video after the lane-changing behavior, if the following vehicle (vehicle E) overtakes the lane-changing vehicle (vehicle A) quickly, or the gaps between them (G_4) decreases quickly, we consider the lane change belongs to this subtype.

X2: The inexplicable ones. In the other N_{X2}=88 cases, the reasons for lane-changing are not clear, and we cannot find any ordinary benefit from these cases. Even after carefully watching the video, it is still too difficult to deduce the reasons. Maybe why the driver changes lane is due to some special personal preference, or they indeed belong to MLCs. However, the validity of these hypotheses are difficult to check. For example, we know there are some drivers who have special preferences, but we don't know how one specific driver makes decisions; we think there should be some MLCs in this highway, but we don't know whether one specific lane change belongs to MLC or not. Therefore, we could only left them as "inexplicable" in this paper, and they need to be investigated in the future.

Besides, when the lane-changing vehicles' velocities are smaller than 10m/s, we find there are some special lane changes. They could only be observed in the dataset of Camera 8 (8:05-8:20), when the densities are very high. In this dataset, the proportion of Group X2 is much higher than that in other datasets. Among them, there are even 3 cases in which $G_2 < 0$.

One typical example is shown in Fig.8. In the yellow circle of Fig.8(a), it is impossible for the vehicle on Lane 5 (Vehicle A) to change lane, since it is hindered by the other one on Lane 4 (Vehicle B). It seems that there is no need to change lane at this moment, since the averaged velocities on Lane 4 and 5 are nearly the same. But Vehicle A still tries to do so, as shown in Fig.8(b). During this process, Vehicle A has to decelerate, rather than acceleration in many other cases. And in Fig.8(c), Vehicle A comes to the back of Vehicle B, but the velocity becomes much slower than before. It seems that Vehicle A has decided to move to the back of Vehicle B at the beginning, but why it wants to do that is not clear.

In short, we think the lane changes when $V_0 < 10m/s$ are quite different from that when$V_0 \geq 10\mathrm{m/s}$, and the rules need to be studied independently. However, it is very difficult to quantitatively determine the rules with current data, since the sample is not large enough: these special behaviors are not observed in other NGSIM video data, including that of I-80. Thus in the future, we still need more empirical data for study.



(a)

(b)

(c)

Fig.8. The special lane changes found at camera 8: the front gap in the target lane is smaller than 0. (a) 0:19; (b) 0:21; (c) 0:23.

In a word, we think the new lane-changing rule can be described as "to have a further position within 9 seconds". When one of the three equations are fulfilled, the vehicles may choose to change lanes:

$$\begin{cases} (\mathrm{G}_2 > \mathrm{G}_1 \text{ and } \mathrm{V}_1 < \mathrm{V}_2) \\ \text{or } (\mathrm{G}_2 > \mathrm{G}_1 \text{ and } \mathrm{V}_1 \geq \mathrm{V}_2 \text{ and } \frac{\mathrm{G}_2 - \mathrm{G}_1}{\mathrm{V}_1 - \mathrm{V}_2} \geq 9\mathrm{s}) \\ \text{or } (\mathrm{G}_2 \leq \mathrm{G}_1 \text{ and } \mathrm{V}_1 < \mathrm{V}_2 \text{ and } \frac{\mathrm{G}_2 - \mathrm{G}_1}{\mathrm{V}_1 - \mathrm{V}_2} < 9s) \end{cases} \quad (7)$$

The proposed model rules could be applied in the velocity range when $10\mathrm{m/s} \leq \mathrm{V_0} \leq 20\mathrm{m/s}$. It can be easily used in the rule-based models, especially the CA models.

The final results of the 373 lane changes are graphically shown in Fig.9. According to Equation (7), we present the proportions of Group A, BC1, X1 and X2, rather than that of Group A, B, C and D. For all the lane-changing data, the $\mathrm{EP_{all}}$ should be:

$$\mathrm{EP_{all}} = \frac{\mathrm{N_A}+\mathrm{N_{BC1}}+\mathrm{N_{X1}}}{\mathrm{N_L}} \quad (8)$$

The result is 30%+43%+3%=76%. It performs much better than the classical rules introduced in Section 3.

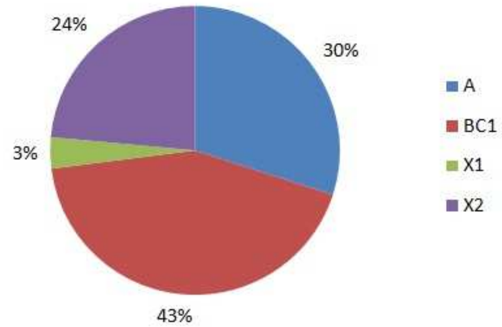


Fig.9. The proportions of the 373 lane changes when $\mathrm{T_c} = 9\mathrm{s}$.

Finally, it is possible to study the lane-changing probability with these data. This important factor could be calculated by:

$$\mathrm{p} = \frac{\mathrm{N_c}}{\mathrm{N_f}} * 100\% \qquad (10)$$

where $\mathrm{N_c}$ = the number of vehicles who fulfill the lane-changing rules and change lanes;

$\mathrm{N_f}$ = the number of vehicles who fulfill the lane-changing rules.

In our data, there is $\mathrm{N_c}$=288, $\mathrm{N_f} \approx 59500$ and $\mathrm{p} \approx 0.5\%$. This means that in each time step 0.5% of the vehicles changes lanes. This value is lower than that used in many previous CA models (e.g., 20% or 30%). But it coincides with some previous empirical data (Knoop et al., 2012), e.g., approximately 0.5 LC/veh/km.

6. The situations when lane changes do not occur

Previous studies of lane-changing usually only comment on the correctness of predicting lane changes, as we did above. However, the correctness of predicting the situations when no lane changes occur in reality is also very important (Knoop and Buisson, 2015). In this paper, we explicitly want to address it. In order to get this, we choose some non-lane-changing vehicles in the same dataset (the video data at Location 1,2,7,8 of US-101 highway). The time interval of collecting data is set as 5 seconds, and then, in each 30-minutes video we can obtain the data of 360 vehicles. The methods of choosing and recording non-lane-changing vehicles are as follows:

(1) All the chosen vehicles are the rightmost ones on the certain lane. This can make sure that the data of the front vehicles on all the lanes can be collected, and the status of the chosen vehicles can be clearly observed in the following several seconds.

(2) All the chosen vehicles do not change lanes in the corresponding video. It is possible that some chosen vehicles change lanes at upstream or downstream locations, but it does not matter. Based on (1), we can make sure that they do not change lanes within at least 5 seconds.

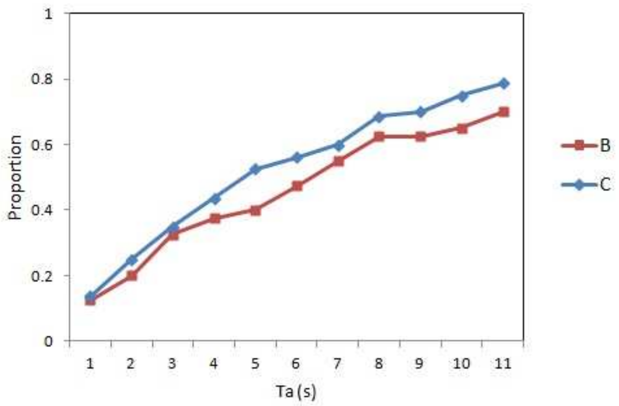
(3) Only small cars are chosen, since large vehicles usually do not change lanes due to their bad driving performance.

(4) The lanes are chosen in turns, and the lane number X is set as X=MOD(T/15, 5)+1. Here T is the time of collecting data.

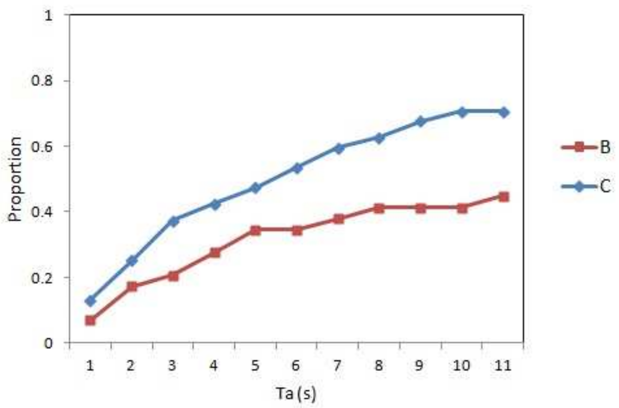
(5) Sometimes the moment of collecting data is slightly changed. For example, at 3:00, if the rightmost vehicle on the certain lane is one truck or it changes lane several seconds later, we will choose the rightmost one on the same lane (the following vehicle) which appears at 3:01 or 3:02.

(6) Sometimes it is impossible for the chosen vehicle to change lanes, since its left or right lanes are partly (or completely) occupied by other vehicles. At this situation, the attempt of lane-changing may immediately lead to traffic accidents. This can frequently occur when the density is high. For this situation, we just consider it as Group "impossible" (Im for short).

(7) If the vehicle is "possible" to change lanes, when the lane number is 2,3 or 4, there are two alternative lanes which can be chosen for the attempts of lane-changing. For all the non-lane-changing vehicles on these 3 lanes, we consider both lanes and calculate two values of T_a. But for the vehicles on Lane 1 or 5, they only have one possible lane and one T_a.



(a)



(b)

Fig.10. The cumulative curve of T_a values in Group B and C of non-lane-changing vehicles. (a) R-L attempts. (b) L-R attempts.

Among the results of the 1440 non-lane-changing vehicles, the Right-to-Left (R-L for short) attempts and that of Left-to-Right (L-R for short) ones are analyzed respectively. For both directions, there are 1152 records, which are close to the number of lane changes. Except the vehicles which belong to Group Im, the others also can be classified into Group A, B, C, D, and the data in Group B, C need to be investigated. Their results are shown in Fig.10 and Fig.11. In Fig.10(a)(b), the tendencies of the cumulative curves in Group B and C are similar. If we compare them with Fig.6, we can say all of them are qualitatively similar.

In Fig.11, the results are also similar. For non-lane-changing vehicles, the equation for $\mathrm{EP_{B+C}}$ is the same as Equation (6). We find the $\mathrm{EP_{B+C}}$ monotonically decrease at both situations. When $9\mathrm{s} \leq \mathrm{T_c} < 11s$, the two curves in Fig.11 also become nearly flat, which is similar to that in Fig.7. Besides, for non-lane-changing vehicles Group X is constituted by Group BC2 and Group D, since there does not exist Group X1.

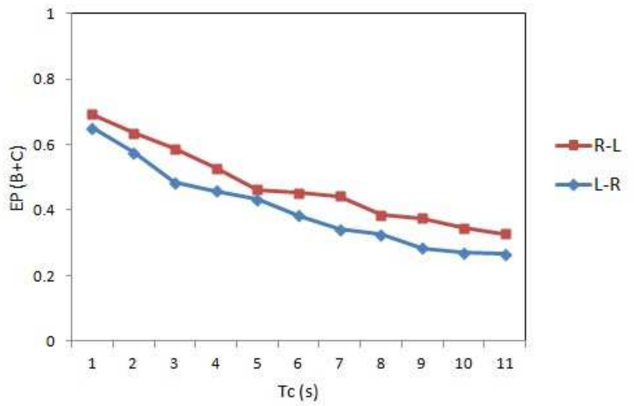


Fig.11. The EPs of new rules in Group B+C of non-lane-changing vehicles.

Then we check whether the lane-change rule, including the time horizon $\mathrm{T_c}$, is also feasible for the non lane-changing cases. The data of L-R attempts and R-L attempts are combined, due to their similar characteristics shown in Fig.10 and 11. Thus the total number of attempts is $\mathrm{N_{NL}}$=2304. Note that the calculation of $\mathrm{EP_{all}}$ of non-lane-changing vehicles is quite different from Equation (8):

$$\mathrm{EP_{all}} = \frac{\mathrm{N_{BC2}+N_D+N_{Im}}}{\mathrm{N_L}} \quad (9)$$

Then the results of $\mathrm{EP_{all}}$ at different situations are shown in Fig.12. When $\mathrm{T_c} < 9s$, the tendencies are the same as that of $\mathrm{EP_{B+C}}$ in Fig.7 and Fig.11. The curve of lane-changing vehicles gradually increase, while that of non-lane-changing vehicles slightly decrease. When $\mathrm{T_c} \geq 9\mathrm{s}$, both curves become quite close and keep nearly constant, which means one steady state is obtained. Here we show the averaged value of both curves, and we find the maximum value also appears at $\mathrm{T_c} = 9\mathrm{s}$. The corresponding values of $\mathrm{EP_{all}}$ for lane-changing vehicles, non-lane-changing vehicles and both are about 76%, 81% and 79%. These results are significantly higher than that of the three classical rules, which further prove the validity of our lane-changing rules.

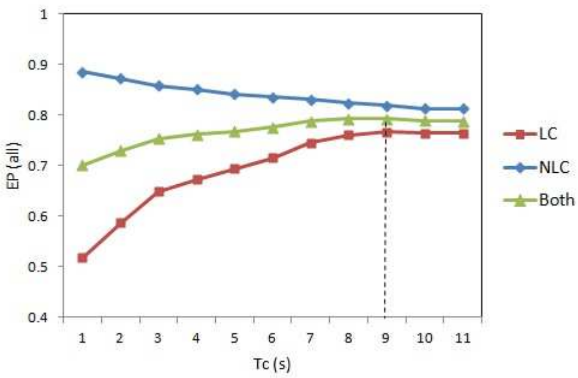
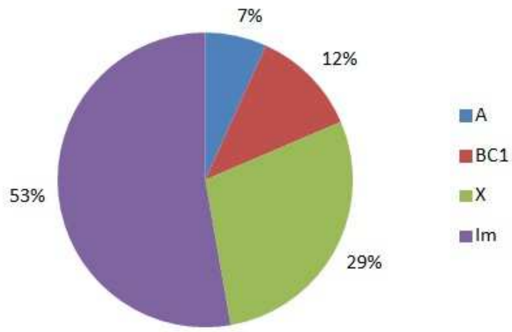
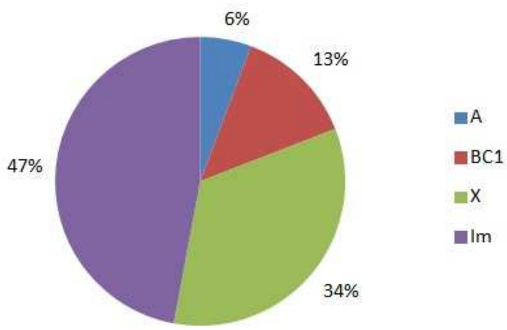


Fig.12. The EPs of new rules in all the lane-changing and non-lane-changing vehicles.

Besides, when $T_c = 9s$ is used for distinction, the proportions of four groups (A, BC1, X and Im) in R-L attempts and L-R attempts are shown in Fig.13. The results are also nearly the same.



(a)



(b)

Fig.13.The proportions of the 1440 non-lane-changing vehicles when $T_c = 9s$. (a) R-L attempts. (b) L-R attempts.

Finally, we would like to discuss the direction of lane changes. We think the two directions should be equally treated in the study, and the reasons are:

(1) As discussed before (Nagel et al., 1998), the rule for lane changing in USA is "symmetric", rather than the "asymmetric" one in Germany. Overtaking on the right lane is also possible in the empirical data of USA. For DLCs, the two directions should be theoretically equal.

(2) In the results of lane-changing vehicles, the difference between two directions is not significant. In all the 373 lane changes, 301 vehicles move to left and 72 vehicles move to right.

For the former, the proportion of Group X2 is 21%, while for the latter the proportion is 33%. Although the result for L-R ones is a little higher, we cannot simply consider all of them as MLCs. Actually, we think there are also some MLCs in R-L ones.

(3) In the results of non-lane-changing vehicles, the features of L-R attempts and the R-L attempts are not only qualitatively, but also quantitatively the same (see Fig.10, 11, 13). On one hand, it means in this dataset, the spatial distribution of vehicles on the five lanes is homogeneous (This is also determined by the symmetric lane-changing rules in USA). On the other hand, it can explain why the L-R lane changes and the R-L ones have similar features, and why they can be modeled by the same rules.

7. Conclusion

In this paper we study the explanatory variables for discretionary lane changes in highway traffic flow. The NGSIM video data are used, and the 373 lane changes at 4 locations in US-101 highway are analyzed and classified. The classical concept of rule-based models which mainly considers the current status of adjacent vehicles, and the new concept which predicts the future movements after some time horizon with or without lane-changes, are both compared with the empirical dataset. We find the classical concepts cannot explain many of the lane changes occurred in these data. On the contrary, the new concept can explain most of them (76%). This new concept can be described as "to have a further position within 9 seconds", which is simple and easy to understand. It is also easy to be used in the microscopic traffic flow simulation, and can form a basis for traffic control which takes lane-changings into account. Besides, we also study the data of some non-lane-changing vehicles, and we find most of them (81%) cannot fulfill the new lane-changing rules. This means the new concept can be considered as one sufficient and necessary condition for DLC. Therefore, we think this work make some contribution to this field.

Nevertheless, there are still many problems to be solved. As we mentioned before, the inexplicable lane changes (24%) in the dataset need to be further studied; the lane-changing rules at low velocities ($V_0 < 10m/s$) and high densities need to be separately investigated; heterogeneous lane-changing model in which different drivers have different properties also needs to be considered in the future, etc. Besides, for the study of DLCs, the limitation of NGSIM data is clear: most sections are not far from ramps and MLCs cannot be simply excluded. Thus more empirical data is needed for further checking, especially the data in some other locations or countries.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (No. 51478113, 51608115), the Natural Science Foundation of Jiangsu Province (Grant No.BK20150613, No.BK20150619), and the Netherlands Organization of Scientific Research (NWO).

References

Ahmed, K.I., Ben-Akiva, M.E., Koutsopoulos, H.N., Mishalani, R.G., 1996. Models of freeway lane

changing and gap acceptance behavior. In: Proceedings of the 13th International Symposium on the Theory of Traffic Flow and Transportation, pp. 501–515.
Chang, G. L., Kao, Y. M, 1991. An empirical investigation of macroscopic lane-changing characteristics on uncongested multilane freeways. Transportation Research Part A, 25, 375-389.
Choudhury, C. F., Ben-Akiva, M. E., 2013. Modelling driving decisions: a latent plan approach. Transportmetrica A, 9, 546-566.
Chowdhury, D., Wolf, D. E., Schreckenberg, M., 1997. Particle hopping models for two-lane traffic with two kinds of vehicles: Effects of lane-changing rules. Physica A, 235, 417-439.
Delpiano, R., Herrera, J.C., Coeymans, J.E., 2015. Characteristics of lateral vehicle interaction, Transportmetrica A, 11, 636-647.
FHWA, 2008. The Next Generation Simulation (NGSIM). <http://www.ngsim.fhwa.dot.gov/>
Gipps, P. G., 1986. A model for the structure of lane-changing decisions. Transportation Research Part B, 20, 403-414.
Hu, X., Wang, W., Yang, H., 2012. Mixed traffic flow model considering illegal lane-changing behavior: Simulations in the framework of Kerner's three-phase theory. Physica A, 391, 5102-5111.
Hunt, J. G. , Lyons, G. D., 1994. Modeling dual carriageway lane-changing using neural networks. Transportation Research Part C, 2, 231–245.
Jia, B., Jiang, R., Wu, Q. S., Hu, M. B., 2005. Honk effect in the two-lane cellular automaton model for traffic flow. Physica A, 348, 544-552.
Jin, W. L., 2010. A kinematic wave theory of lane-changing traffic flow. Transportation research part B, 44, 1001-1021.
Kesting, A., Treiber, M., and Helbing, D., 2007. General lane-changing model MOBIL for car-following models. Transportation Research Record, 1999, 86–94.
Kukida, S., Tanimoto, J., Hagishima, A., 2011. Analysis of the influence of lane changing on traffic-flow dynamics based on the cellular automaton model. International Journal of Modern Physics C, 22, 271-281.
Knospe, W., Santen, L., Schadschneider, A., Schreckenberg, M., 1999. Disorder effects in cellular automata for two-lane traffic. Physica A, 265, 614-633.
Knospe, W., Santen, L., Schadschneider, A., Schreckenberg, M., 2002. A realistic two-lane traffic model for highway traffic. Journal of Physics A, 35, 3369.
Knoop, V.L., Hoogendoorn, S.P., Siomi, Y., Buisson, Ch., 2012. Quantifying the Number of Lane Changes in Traffic: Empirical Analysis. Transportation Research Records 2278, 31-41.
Knoop, V.L., Buisson, Ch., 2015. Calibration and Validation of Probabilistic Discretionary Lane-Change Models. IEEE Transactions on Intelligent Transportation Systems 16, 834 - 843.
Laval, J. A., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transportation Research Part B, 42, 511-522.
Lee, J., Park, M., Yeo, H., 2016. A Probability Model for Discretionary Lane Changes in Highways. KSCE Journal of Civil Engineering, 20, 2938-2946.
Leclercq, L., Chiabaut, N., Laval, J., Buisson, C., 2007. Relaxation phenomenon after lane changing: experimental validation with NGSIM data set. Transportation Research Record, 1999, 79-85.
Li, X. G., Jia, B., Gao, Z. Y., Jiang, R., 2006. A realistic two-lane cellular automata traffic model considering aggressive lane-changing behavior of fast vehicle. Physica A, 367, 479-486.
Moridpour, S., Rose, G., Sarvi, M., 2010. Effect of surrounding traffic characteristics on lane

changing behavior. Journal of Transportation Engineering, 136, 973-985.
Moridpour, S., Sarvi, M., Rose, G., 2013. Lane changing models: a critical review. Transportation letters.
Nagel, K., Schreckenberg, M., 1992. A cellular automaton model for freeway traffic. Journal de physique I, 2, 2221-2229.
Nagel, K., Wolf, D. E., Wagner, P., Simon, P., 1998. Two-lane traffic rules for cellular automata: A systematic approach. Physical Review E, 58, 1425.
Park, M., Jang, K., Lee, J., Yeo, H., 2015. Logistic regression model for discretionary lane changing under congested traffic. Transportmetrica A, 11, 333–344.
Rahman, M., Chowdhury, M., Xie, Y., He, Y., 2013. Review of microscopic lane-changing models and future research opportunities. IEEE transactions on intelligent transportation systems, 14, 1942-1956.
Schakel, W.J., Knoop, V.L., Van Arem, B., 2012. LMRS: Integrated Lane Change Model with Relaxation and Synchronization, Transportation Research Record, 2316, 47-57.
Sheu, J. B., Ritchie, S. G., 2001. Stochastic modeling and real-time prediction of vehicular lane-changing behavior. Transportation Research Part B, 35, 695-716.
Sheu, J. B., 2013. Microscopic traffic behaviour modelling and simulation for lane-blocking arterial incidents. Transportmetrica A, 9, 335-357.
Sun, D., Elefteriadou, L., 2010. Research and implementation of lane-changing model based on driver behavior. Transportation Research Record, 2161, 1-10.
Sun, D. J., Kondyli, A., 2010. Modeling Vehicle Interactions during Lane-Changing Behavior on Arterial Streets. Computer-Aided Civil and Infrastructure Engineering, 25, 557-571.
Thiemann, C., Treiber, M., Kesting, A., 2008. Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data. Transportation Research Record, 2088, 90-101.
Toledo, T., Koutsopoulos, H. N., Ben-Akiva, M., 2007. Integrated driving behavior modeling. Transportation Research Part C, 15, 96-112.
Wagner, P., Nagel, K., Wolf, D. E., 1997. Realistic multi-lane traffic rules for cellular automata. Physica A, 234, 687-698.
Wang, Q., Li, Z., Li, L., 2014. Investigation of Discretionary Lane-Change Characteristics Using Next-Generation Simulation Data Sets. Journal of Intelligent Transportation Systems, 18, 246–253.
Wolfram, S., 1983. Statistical mechanics of cellular automata. Reviews of modern physics, 55, 601.
Wu, J., Brackstone, M., McDonald, M., 2003. The validation of a microscopic simulation model: a methodological case study. Transportation Research Part C, 11, 463-479.
Yang, Q., Koutsopoulos, H.N., 1996. A microscopic traffic simulator for evaluation of dynamic traffic management systems. Transportation Research Part C, 4, 113–129.
Yeo, H., Skabardonis, A., 2008. Parameter estimation for NGSIM freeway flow algorithm, ASCE AATT 2008.
Yeo, H., et al., 2008. Oversaturated freeway flow algorithm for use in next generation simulation. Transportation Research Record, 2088, 68-79.
Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics. Transportation research part C, 26, 367-379.

Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. Transportation research part B, 60, 16-32.