A causal inference based speed control framework for discretionary lane-changing process

Zhen Zhou
College of Automotive and Traffic Engineering
Nanjing Forestry University, Nanjing, P.R. China, 210037
and
Department of Statistics
University of Illinois Urbana-Champaign, Champaign, Illinois, The United States, 61820
Email: zhenz5@illinois.edu

Yi Zhao*
College of Automotive and Traffic Engineering
Nanjing Forestry University, Nanjing, P.R. China, 210037
Email: zhaoyi207@njfu.edu.cn

Minghao Li
College of Automobile and Traffic Engineering
Nanjing Forestry University, Nanjing, P.R. China, 210037
Email: limh@njfu.edu.cn

Yuyang Bao
College of Science
Wuhan University of Technology, Wuhan, P.R. China ,430070
Email: byy123456@whut.edu.cn

*= Corresponding Author
Word Count: 4798 words + 3 table (250 words per table) = 5548 words

Submitted [07/31/2022]

ABSTRACT

This research proposes a causal inference based framework to help improve the safety, efficiency and eco-friendly of lane change behavior. The framework can be divided into three parts. Firstly, lane change data is used to calculate causal effect between generated variables and the lane change duration. The second part is the calculation of causal average treatment effect by applying causal forest which takes treatment variable, outcome variable and convarites into consideration. The research also constructs an optimization model to help find the optimal velocity and acceleration while using time-to-collision (TTC) to ensure the safety of lane change. Numeric experiment is also conducted to validate the performance of the framework. Although the correlation between variables and duration is weak, the causal relationship can be explored by calculating the treatment effect. And an optimal velocity and acceleration control scheme is formed by using the framework, which can reduce the lane change duration at the same time. The result shows the framework can improve the efficiency and eco-friendly of the lane change behavior while ensuring TTC in the safe interval.

Keywords: Causal inference, Lane change behavior, Treatment effect, Speed control framework

INTRODUCTION

Vehicle lane changing is the main traffic behavior that triggers traffic flow fluctuations and traffic accidents. So researchers have done a lot of research on vehicle lane changing behavior, including the study of lane change behavior characteristics in terms of acceptable lane change gap, lane change duration [1]and lane change probability[2], the study of lane change behavior modeling based on game theory[3] and deep learning[4], the study of lane change behavior for safety risk assessment and lane change warning[5], and also includes the study of lane change trajectory planning and lane change behavior impact. With the arrival of Cooperative Vehicle Infrastructure System (CVIS) and driverless technologies, the characteristics, decisions and impacts of lane-changing behaviors in mixed traffic flow or driverless environments have gradually become a hot topic of current research.

The connected environment provides information via vehicle-to vehicle and vehicle-to-infrastructure communications. Yasir Ali[6] expressed their understanding on the discretionary lane-changing behaviour in the connected environment, and found that the connected environment has a significant impact on spacing, speed, acceleration noise, and others. They also proposed a game theory-based approach for modelling mandatory lane-changing behaviour in a connected environment in other researches[7]. And a hazard-based duration model was built to quantify the impact of connected driving environment on safety during mandatory lane-changing[8]. Other studies in lane-changing under the connected environment include researches from Lin Dianchao (2019) [9], Talebpour Alireza (2015)[10], Lyu Nengchao (2019)[11], Cao Peng (2017)[12], Mai Trung (2016)[13], Luo Yugong (2016)[14], Desiraju Divya (2015)[15] et al.

The connected environment has shown promise in solving massive traffic difficulties related to safety, efficience, and the environmental impact. But a fully driverless environment or a high degree of CVIS is still a long way off before full-scale promotion. Before this, Assisted driving will be the best way to take full advantage of the traffic information provided by the connected environment. This paper will focus on the vehicle lane change behavior. The lane change behavior can be divided into three phases, which are lane change initiation phase, lane change progress phase and lane change completion phase, the three phases are shown in **Figure 1**. The interaction between the lane change vehicle and the surrounding vehicles provides rich information on traffic and vehicle control. The surrounding vehicles can be simplified as five vehicles, which are front vehicle in current lane, rear vehicle in current lane, front vehicle in target lane, rear vehicle in target lane and overlap vehicle in target lane. In the process of vehicle lane change, how to use the influence relationship between vehicles and accurate information of the surrounding vehicles provided by connected environment? How to integrated safety, efficiency, and the environmental impact needs for scientific control of lane change vehicles? These two problems are the subject of this study. The scientific problem of this study is to explore the impact of vehicle driving state on safety, efficiency, and the environment.

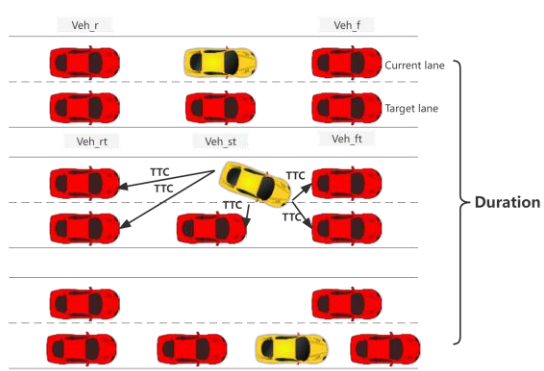

Figure 1 Lane change behavior

In the year of 2021, the Nobel prize in economics goes to David Card, Josh Angrist, and Guido Imbens for their work in the field of causal inference. Causal inference, which is different from correlation analysis, can be used to measure the causality, and it means that causal inference can infer whether a change in one event is the cause of a change in other events [16]. The concept of causal effect was originally introduced by Neyman in 1923 [17], and the framework of causal inference was first proposed by Paul W. Holland in 1986 [18] which is known as Rubin Causal Model (RCM). Another popular causal model is causal diagram model, which was proposed by Pearl [19] (1995). Causal diagram model can be divided into three steps, which are causal discovery, causal effect and refuted estimate. It can be proven that the causal diagram model is equivalent to RCM [20]. Due to the ease of interpretation of causal diagram model, causal it is more often used in other disciplines, such as social science, computer science, and a range of data-related disciplines.

Causal inference is also widly applied in transportation. For example, Li [21] et al. (2020) converts risk identification to the causal effect task and "present a novel two-stage risk object identification framework based on causal inference with the proposed object-level manipulable driving model". You and Han [22] (2020) use causal inference to compute causality in the different frames of traffic accident videos. Cao [23] et al. (2021) proposed a causal based framework to determine impacted region and quantify speed reduction. Ali [24] et al. (2019) use the causal inference model to mine and analyze the causes of traffic safety accidents, including urban road class, urban characteristics, population density, education level, etc. Yang [25] et al. (2021) identified different street impact coefficients on the road system and calculate the causal interaction effects between different streets. Molavipour [26] et al. (2020) explore the interaction of traffic flows at different nodes in the traffic network, which is based entirely on theoretical proof and aim to construct cause-and-effect diagrams for traffic flow.

As an important method for discovering causality between events, it is proposed to use causal inference to explore the relationship between driving behavior control and operational efficiency, traffic safety and environment during lane change in this paper. And a causal inference based speed control system should be proposed to improve discretionary lane-changing process. This system is expected to be applied to vehicle assisted driving systems in the future.

The rest of the article is structured as follows. The METHODS is divided into three sections. The first section introduces the estimation of treatment effects and refutation, the second section introduces the estimation of conditional average treatment effect, and the third section builds a decision system. The DATA DESCRIPTION decribles the basic information of data and correlation analysis is conducted. CASE STUDY is also divided into three parts, which respectively show the results of causal effect analysis, conditional average treatment effect analysis and lane change decision system. Finally, the CONCLUSION section summarizes the conclusions and achievements of this study, and proposes the directions for future research.

METHODS

Estimation of treatment effects and refutation

Causal inference can help us discover the causality between variables, and it can also enhance the interpretability of the model. Treatment effect is an important tool which can quantify the causal effects between variables, it can be estimated from treatment variable and outcome variable.

According to causal inference, if the treatment variable is binary variable, the average treatment effect is

$$E[E(Y|W,T=1)]-E[E(Y|W,T=0)] \quad (1)$$

Where T is treatment variable which is chosen from variable set W, Y is outcome variable. According to **Equation (1)**, we can estimate $E(Y|W,T)$ by using regression model, which can be denoted as $E(Y|W,T)=\beta_0+\beta_t T+\beta_w W$. The average treatment effect is β_t.

In order to test the validation of the estimator, placebo is added to refit the outcome model. We can calculate new effects and compare new effects with the estimated effects. By definition, "a placebo is any treatment that has no active properties" [27], if we replace our treatment variables with placebo, new treatment effects with different distribution can occur. By using this property, we can easily determine whether there is a treatment effect of the original variable.

We can obtaine guidelines for determining whether there is a causal effect between the variable and the outcome variable. Firstly, the estimated effect should not be zero, and the value of estimated effect is proportional to the causal effect of the treatment variable on the outcome variable. Secondly, the placebo effect should go to zero. For box plots, the average value must be 0, and the data in the box plot needs to be significantly different from the estimated value.

Estimation of conditional average treatment effect

If we want to determine the heterogeneity between variables, conditional average treatment effect is a tool that can detect heterogeneity between variables. The input variables of the estimation of conditional average treatment effect are treatment variable, outcome variable and covariates.

We first assume that the outcome variable and treatment variables are continuous variables, tree model is applied to get causal effect of treatment variables. Conditional regression function can be denoted as $\mu_t(X) = E[Y(t)|X]$, while $\hat{\mu}_t(X) = \mu_t(X, \widehat{W}_t) = E[Y(t)|X, \widehat{W}_t]$ is the estimated outcome regression, X are covariates, T is treatment variable which is chosen from variable set W and $Y(t)$ is outcome variable under treatment $T = t$.

In order to calculate treatment effects, in this section, gradient boosting regression is first applied to fit $\mu(X, W) = E[Y|X, W]$ and $f(X, W) = E[T|X, W]$, and local estimates $\widehat{\mu_i}(X, W)$ and $\hat{f}(X, W)$ can be obtained. Then, causal forest proposed by Athey [28] (2019) is fitted to solve the “local moment equation problem” [29].

$$E[(Y - E[Y|X, W] - \theta(x)(T - E[T|X, W]))(T - E[T|X, W])|X = x] = 0 \quad (2)$$

In this section, we apply double machine learning version of causal forests to estimate conditional average treatment effects $\hat{\theta}(x)$, according to Athey [28] (2019) and Keith [29] et al. (2019), the estimation is

$$\hat{\theta}(x) = argmin_{\theta} \sum_{i=1}^{n} K_x(X_i) \cdot (Y_i - \widehat{\mu_i}(X_i, W_i) - \theta(T_i - \hat{f}(X_i, W_i)))^2 \quad (3)$$

Where $K_x(X_i)$ is similarity metric. We finally obtain treatment effect estimator by using double machine learning version of causal forests, this estimator can be denoted as $CATE(X, Y)$.

Decision system

Efficiency, safety and eco-friendly are three key objectives of traffic management and control, and also objectives of system optimization in the context of cooperative vehicle infrastructure system (CVIS). In the next decision system to be built for the lane change scenario, the lane change duration is selected to denote efficiency, TTC to denote safety and acceleration to denote eco-friendly. The shorter the lane change duration means the more efficient the lane change is, and the shorter impact of the lane change behavior on other vehicles. The TTC needs to be controlled within a safe range during the lane change to reduce the risk of vehicle collisions. From the aspect of energy consumption and emissions, the more stable the speed, the less energy and exhaust emissions will be consumed. So the system should avoid rapid acceleration or deceleration.

Based on the above considerations, a framework is constructed to help drivers change lanes more safely, efficiently and eco-friendly. The framework is shown in **Figure 2**.

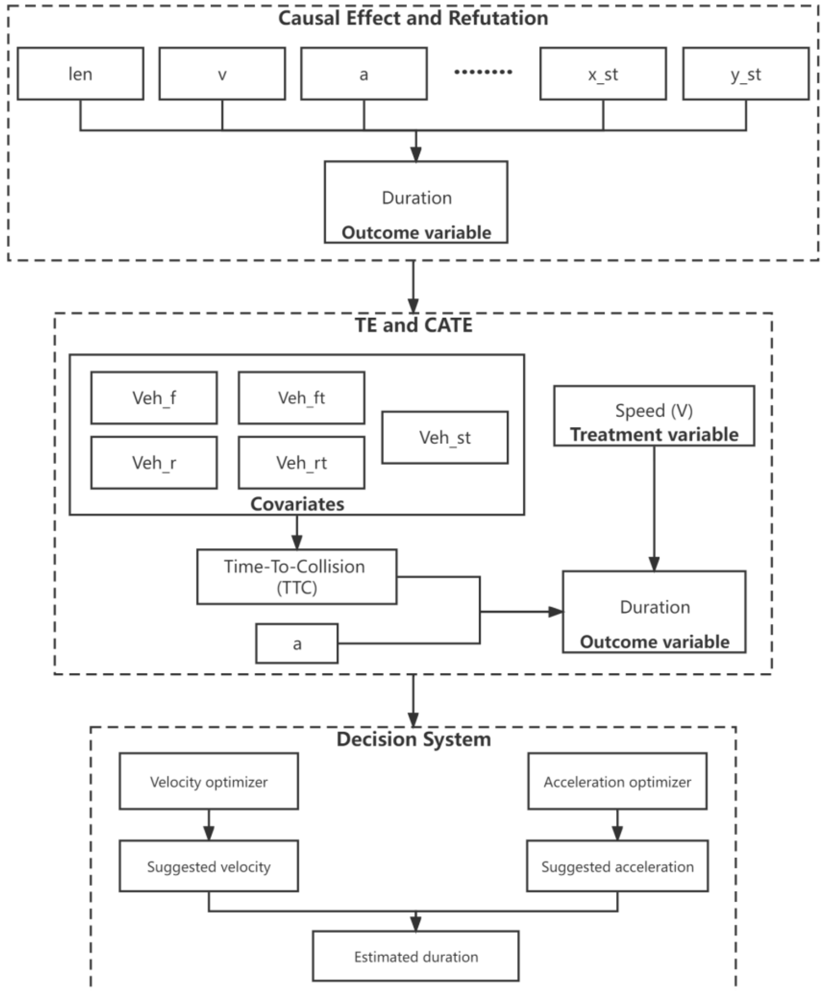

Figure 2 A framework for lane change based causal inference

The conditional average treatment effect is used to measure the "elasticity" of speed with respect to duration, and TTCs are added to the framework to constrain speed. To meet safety, TTCs must be within a reasonable interval *B*, which means $TTCs \in B$, and in this paper, the threshold for TTCs is 2.5 seconds.

The objective function shows in **Equation (4)**.

$$min_v CATE(TTC) + \lambda_1 F(TTC) \tag{4}$$

Where $\lambda \geq 0$, and $F(TTC) = \sum_i TTC_i$. The treatment effect under different TTC can be denoted as $CATE(TTC) = CATE(TTC_{ft}, TTC_f,\ TTC_{rt}, TTC_r,\ TTC_{st})$, it can be estimated by using causal forest mentioned in **Equation (3).** The second term in **Equation (4)** aims to constraint the speed so that it does not choose a high CATE but low speed.

$$min_{h_j} CATE\left(h'_j\right) + \lambda_j \left|h'_j - h_j\right| \tag{5}$$

In order to avoid high energy consumption and high emissions. Acceleration is added into the algorithm. We denoted acceleration as $h \in H$ in this section, and H is the convarites set, which is to prepare for a more general model in the future. From the perspective of efficiency and energy consumption, we use objective function shown in **Equation (5)** to help search the best suggested convaritesn set $\widehat{H}$. The first term in **Equation (5)** is to improve the efficiency of v towards duration under h'_j. The second term aims to help the vehicle transition more smoothly. A decreasing acceleration leads to a decrease in efficiency and a lifting acceleration leads to a high energy consumption and high emissions.

Finally, the whole algorithm for lane change decision system is shown in **Algorithm 1**.

Algorithm 1 Lane change decision algorithm

Input: Training set $D = (x_{ft}, y_{ft}, x_f, y_f, x_{rt}, y_{rt}, y_r, x_{st}, y_{st})$, $V = (v, v_{ft}, v_f, v_{rt}, v_r)$, $duration$; Test set $D' \in \mathbb{R}_{n\times 6}$, $V' = (v'_{ft}, v'_f, v'_{rt}, v'_r, v'_{st}) \in \mathbb{R}_{n\times 5}$; Control variable set H; Coefficients λ

1: Compute $TTC(D, V) = (TTC_{ft}, TTC_f, TTC_{rt}, TTC_r, TTC_{st})$;
2: Compute $CATE(TTC(D, V), treatment = v, outcome = duration)$;
3: **for** D'_i, V'_i in D', V' **do**
4: Generate evenly spaced samples V' and H' according to V and H;
5: Compute $CATE(TTC(D', V', v'), treatment = v', outcome = duration)$;
6: Search v'_i to find $\hat{v}_i$ to minimize $f(v'_i) = CATE(TTC) + \lambda_1 F(TTC)$, such that $TTC(D', V', \hat{v}_i) \in B$;
7: **for** h'_j, h_j in H', H, $j > 1$ **do**
8: Compute $CATE(h'_j, treatment = v', outcome = duration)$;
9: Search h'_j to find $\hat{h}_j$ to minimize $g(h'_j) = CATE(h'_j) + \lambda_j |h'_j - h_j|$;
10: **end for**
11: **end for**
12: $\hat{H} = (\hat{h}_1, \hat{h}_2, ..., \hat{h}_n)$

Output: $\hat{V}$, $\hat{H}$, $CATE$

In the algorithm, line 4 is a function to generate evenly spaced array whose interval is determined by the range of *T*. By learning the training set, we can obtain the treatment effect of velocity for duration at specific TTC. We then solve two types of optimization problems to find the optimal solution for the lane change state. We can use this framework to help drivers change lanes more efficiently while ensuring safe.

DATA DESCRIPTION

The data we use to conduct case study is from previous work which is proposed bt Dong [30] et al., (2017), and NGSIM [31] (2004) is used to extract the lane change behavior. In this section, we add different types of variables which are generated by lane change behavior to the data set. The description of variables is shown in **Table 1.**

TABLE 1 Variable Description

Variables	Description	Variables	Description
len	Vehicle length	f_len	Front vehicle's length in current lane
v	Vehicle velocity	r_len	Rear vehicle's length in current lane
a	Vehicle acceleration	ft_len	Front vehicle's length in target lane
v_after	Vehicle velocity after changing lane	rt_len	Rear vehicle's length in target lane
a_after	Vehicle acceleration after changing lane	st_len	Overlap vehicle's length in target lane
f_v	Front vehicle's velocity in current lane	x_ft	The x distance with front vehicle in target lane
r_v	Rear vehicle's velocity in current lane	y_ft	The y distance with front vehicle in target lane
ft_v	Front vehicle's velocity in target lane	x_rt	The x distance with rear vehicle in target lane
rt_v	Rear vehicle's velocity in target lane	y_rt	The y distance with rear vehicle in target lane
st_v	Overlap vehicle's velocity in target lane	x_f	The x distance with front vehicle in current lane

f_a	Front vehicle's acceleration in current lane	y_f	The y distance with front vehicle in current lane
r_a	Rear vehicle's acceleration in current lane	x_r	The x distance with rear vehicle in current lane
ft_a	Front vehicle's acceleration in target lane	y_r	The y distance with rear vehicle in current lane
rt_a	Rear vehicle's acceleration in target lane	x_st	The x distance with overlap vehicle in target lane
st_a	Overlap vehicle's acceleration in target lane	y_st	The y distance with overlap vehicle in target lane
duration	The duration of lane changing		

There are 544 observed lane change vehicles and 31 variables for each lane change, each observation has a specific lane change trajectory. **Table 2** shows the linear correlation of duration with other variables. It can be found that the biggest absolute linear correlation value is -0.183 obtained by acceleration and duration. All variables have little linear correlation. In the real world, when lane change behavior occurs, most variables must have a strong association with duration. For example, there is definitely a causal relationship between velocity and duration. So, in the next section, we will use causal effects to quantify the relationship between variables and duration.

TABLE 2 The Correlation Between Each Variable and Duration

Variable	Value	Variable	Value	Variable	Value	Variable	Value
len	-0.021	ft_v	-0.131	st_a	-0.088	y_ft	0.068
v	-0.152	rt_v	-0.127	f_len	0.0008	x_rt	0.015
a	-0.089	st_v	-0.099	r_len	-0.015	y_rt	-0.065
v_after	-0.183	f_a	0.014	ft_len	0.074	x_f	0.093
a_after	0.006	r_a	0.028	rt_len	-0.055	y_f	-0.100
f_v	-0.098	ft_a	0.046	st_len	-0.070	x_r	0.049
r_v	-0.114	rt_a	-0.008	x_ft	-0.025	y_r	-0.072

CASE STUDY

Causal Effect Analysis

Causal diagram is a visualization tool which can visualize the causal relationship within variables, and it is a directed acyclic graph (DAG) which is composed of nodes and arrows. Nodes are composed of the treatment variable *T*, the outcome variable *Y* and other variables *W*. The direction of the arrow is the direction of our hypothetical cause relationship. **Figure 3** is the causal diagram which shows the causal relationship of variables in **Table 1**.

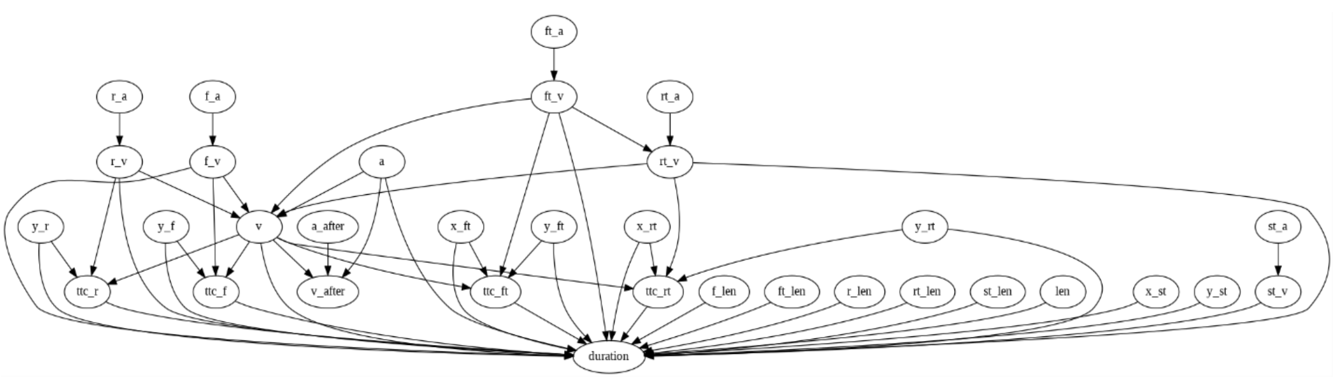

Figure 3 Hypothetical causal relationship between variables

In this section, the average treatment effect of potential treatment variables is calculated, but not include the TTC variable. Python package “Dowhy” [32] (2020) is used to construct a causal diagram and calculate average treatment effect. According to the diagram, outcome variable is duration, and treatment variables show in **Figure 4**.

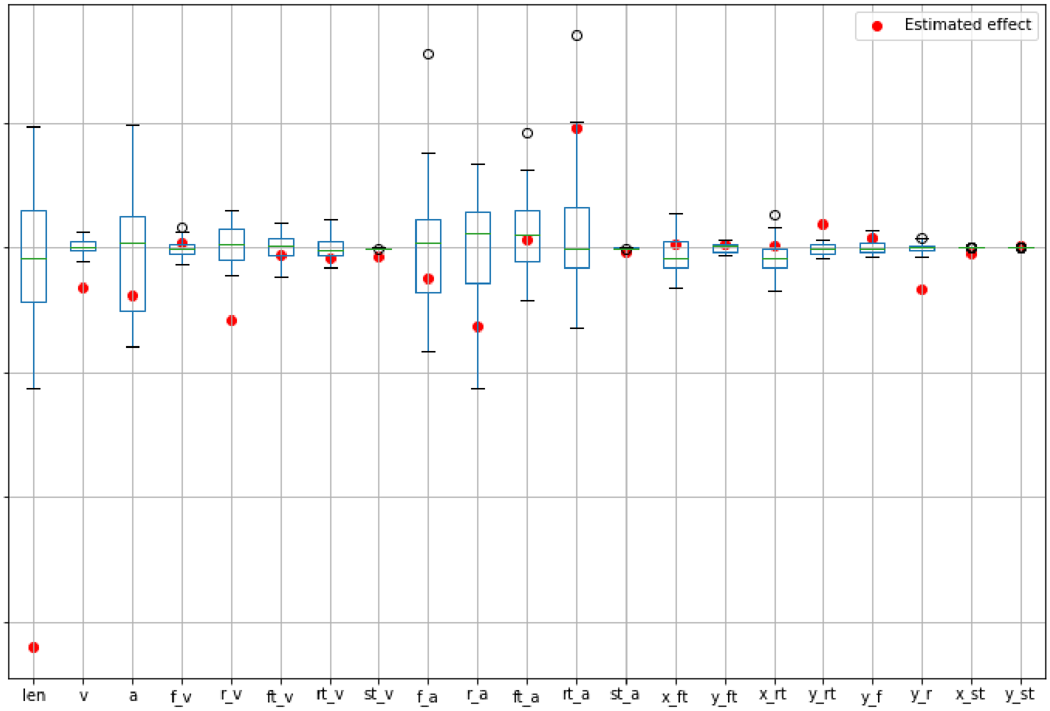

Figure 4 Boxplot of Causal Effects Using Placebo

Back-door criterion is adopted as the identification criteria, and regression is applied to estimate causal effect. In the last step, placebo treatment as refutation method is applied to obtain new treatment effect. Fnally the results of estimated treatment effects are shown in **Figure 4**.

Red dots are estimated effect of different variables. The boxplots are the treatment effects which are generated by 30 randomized placebo tests.

According to the rules of placebo refutation test in the causal inference, the causal effect between each variable and the outcome variable can checked. From **Figure 4**, the estimated effect of “len”, “v”, “r_v”, “st_v”, “st_a”, “y_rt”, “x_st” and “y_r” have a significant difference with the effect using placebo treatment while the effect using placebo treatment tends to be zero. There is a causal effect of these variables with the outcome variable, which is duration.

Classifying the variables according to the level of difference reveals that “len”, “v”, “r_v”, “y_rt”and “y_r” have a relatively higher causal effect on outcome variable; “st_v”, “x_st” and “st_a” have a relatively low causal effect on outcome variable.

By using the positive and negative causal effects for identification we can find that “len”, “v”, “r_v”, “st_v”, “st_a”, “x_st” and “y_r” have a negative causal effect on outcome variable, while “y_rt” has a positive causal effect on outcome variable.

Conditional average treatment effect under TTCs and acceleration

In this section, the double machine learning version of causal forests was used to estimate conditional average treatment effects under TTCs and acceleration. Api from “EconML” (Keith et al., 2019) was used to calculate conditional average treatment effects. There are five parameters in this step, which are “TTC_ft”, “TTC_f”, “TTC_rt”, “TTC_r” and “TTC_st”. These five variables are covariates *X* which have been mentioned in **Equation (2)**. The velocity is treatment variable while duration is outcome variable.

The calculation of different covariates was divided into four parts, front vehicles, rear vehicles, overlap vehicle in target lane and acceleration. Two 3D surface plots and two 2D plot are constructed to show the conditional average treatment effects under different TTCs and acceleration. The results are shown in **Figure** 5.

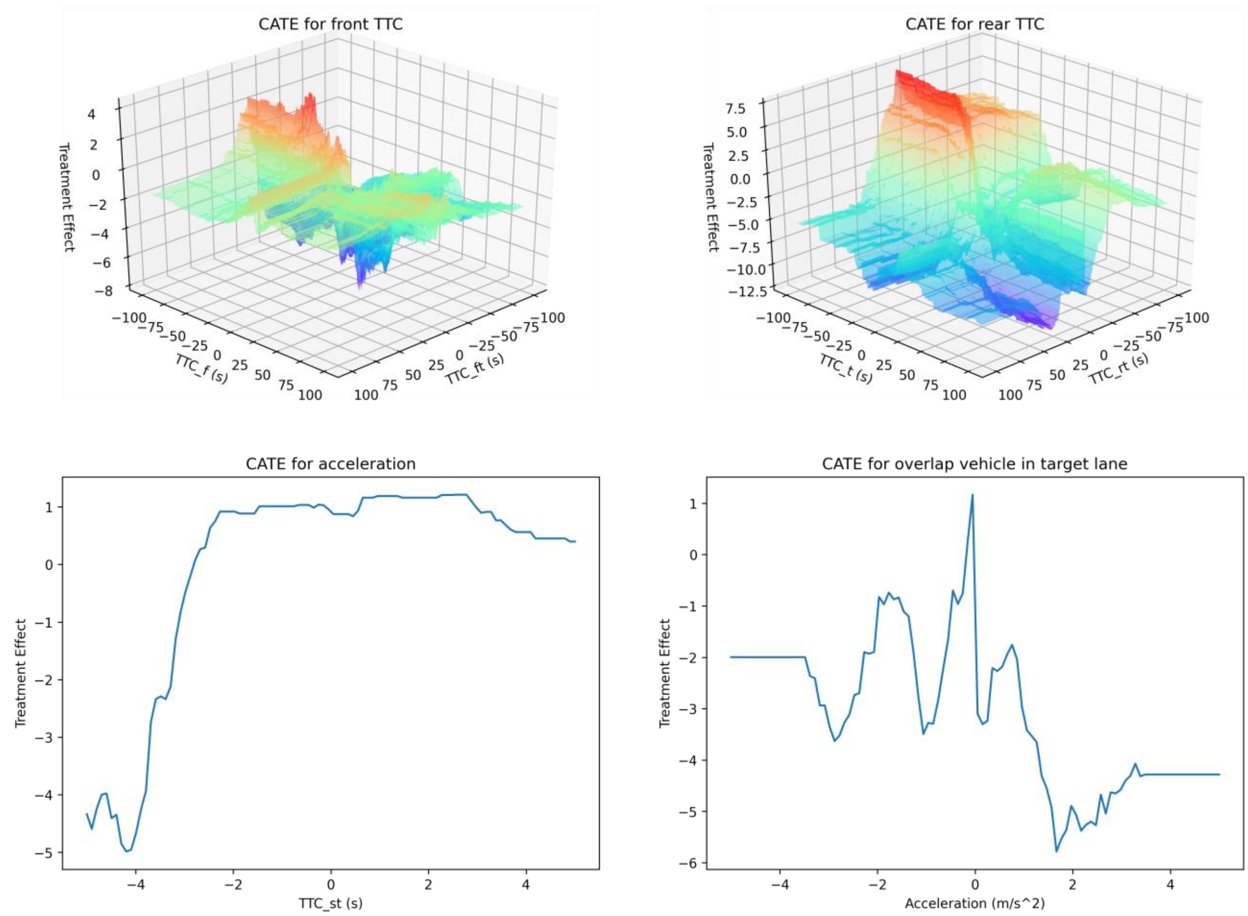

Figure 5 Treatment effects for front, rear vehicles, overlap vehicles and acceleration

Conditional average treatment effects can be used to clarify the heterogenous of treatment effect. Different TTCs or acceleration affect the treatment effect differently. The positive value in the 3D and 2D plot represents the positive “elasticity” of velocity with respect to duration at specific TTCs or acceleration, implying that an increase in velocity causes an increase in duration, which means speed has a positive causal effect on duration. Negative values represent a negative causal effect of speed to duration at specific TTCs or acceleration.

In order to visualize the treatment effect of specific vehicle at different velocities with respect to duration, we add solid black lines which denotes one single car under different velocities. We take two vehicles as examples and take TTCs of rear and front vehicles as covarites to simulate changes, the results were shown in **Figure 6**.

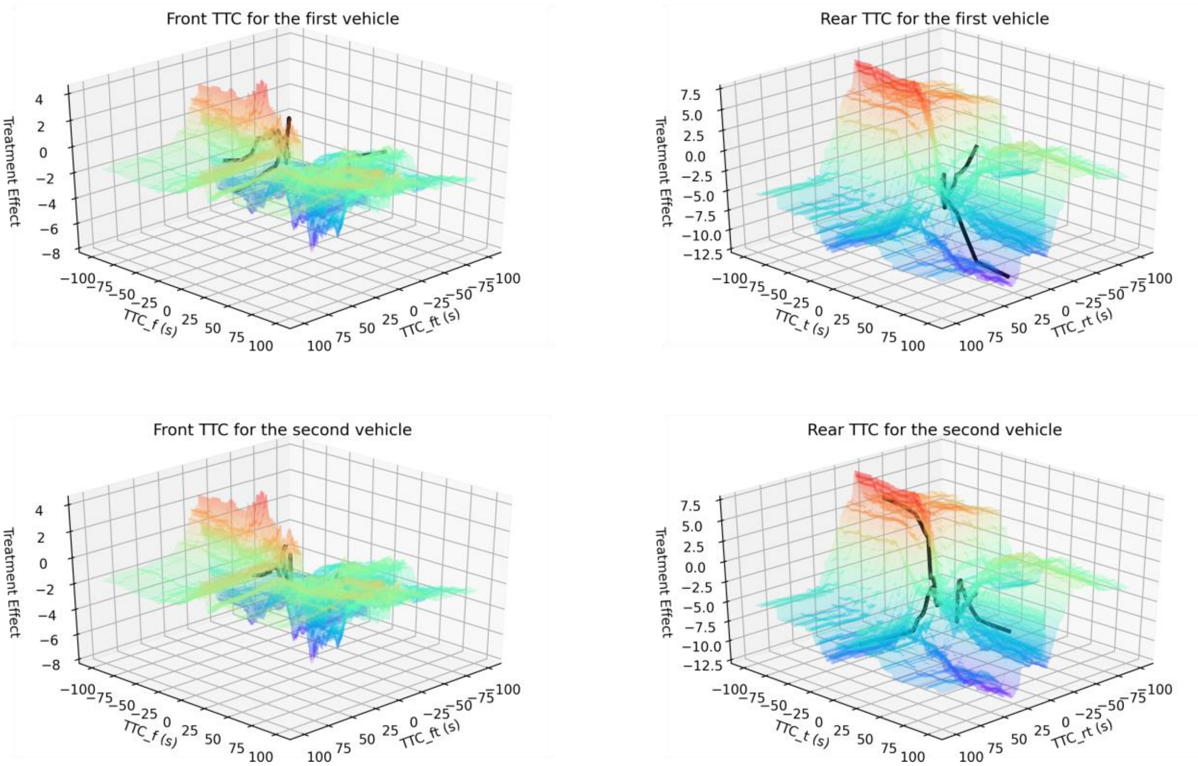

Figure 6 Surface plots of treatment effects for front and rear vehicles under real scenarios

Figure 6 shows the variation of conditional average treatment effect of vehicle speed under different TTC. For different cars, the trajectories of TTCs are different. It can be noticed that not all vehicles can reach the lowest point in surface plot, it depends on the location and velocity of vehicles. Considering the safe range of TTC and the causal effect of velocity on duration, a lower conditional average treatment effect does not mean a better lane change decision. When velocity is low, even if the conditional average treatment effect is very low, the efficiency of the lane change can be very poor. So, we need to take the value of velocity into consideration, and this is the purpose of the second term in **Equation (4)**. For acceleration, considering the efficiency and energy consumption, a lower conditional average treatment effect also does not mean a better lane change decision. The change of acceleration should also be taken into consideration.

Lane change decision system

According to Algorithm 1, after obtaining TTC and treatment effect, the next step is to search velocity to minimize the **Equation (4)** and search covariates T to minimize the **Equation (5)** by using grid search. The suggested velocity and acceleration are calculated to balance the safety and efficiency.

Firstly, the estimated duration of the lane change under different λ_1 and λ_2 should be calculated, the result was shown in **Figure 7**. Surface plot is applied to show the different duration.

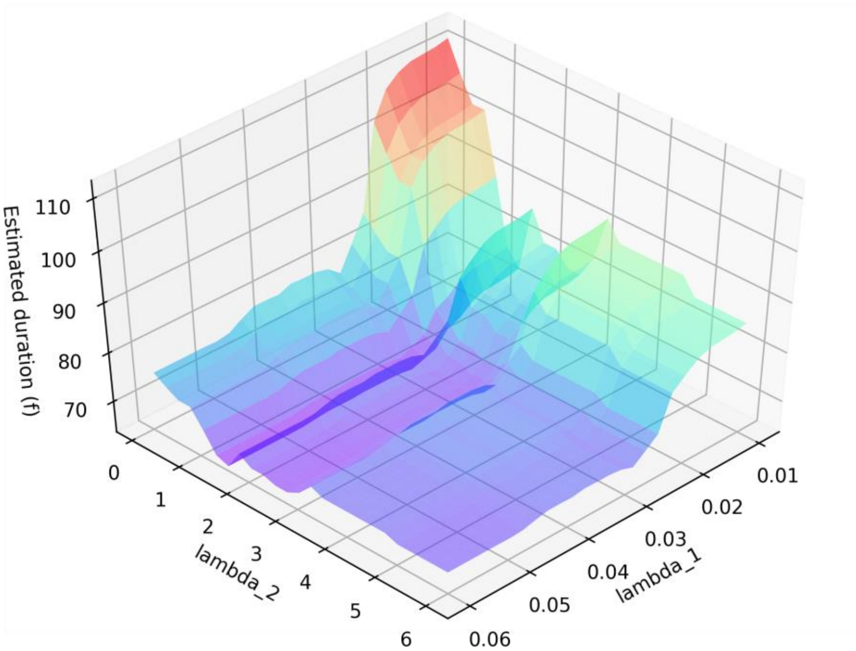

Figure 7 Estimated duration at different λ_1 and λ_2

According to Figure 7, the lowest estimated duration can be obtained when $\lambda_2 = 2\ or\ 3$. Several experiments were implemented to calculate the average velocity, average acceleration, conditional average treatment effect (CATE) and average duration before and after using the framework. The values of parameters are $\lambda_1 = 0.1, 0.2, 0.3, 0.4, 0.5$, and $\lambda_2 = 2, 3$.

TABLE 3 Change in velocity, conditional average treatment effect and average duration before and after using framework

		Without using framework				**Using framework**			
λ_1	λ_2	$\bar{v}$	$\bar{a}$	**CATE**	**Average Duration**	$\bar{v}$	$\bar{a}$	**CATE**	**Average Duration**
0.01	2	7.165	0.196	−2.029	72.043	7.819	0.664	−5.926	82.941
0.01	3	7.165	0.196	−2.029	72.043	7.819	0.545	−5.633	89.189
0.02	2	7.165	0.196	−2.029	72.043	7.797	0.664	−5.924	79.091
0.02	3	7.165	0.196	−2.029	72.043	7.797	0.545	−5.638	77.991
0.03	2	7.165	0.196	−2.029	72.043	7.791	0.664	−5.926	68.973
0.03	3	7.165	0.196	−2.029	72.043	7.791	0.545	−5.633	67.792
0.04	2	7.165	0.196	−2.029	72.043	7.777	0.664	−5.931	68.981
0.04	3	7.165	0.196	−2.029	72.043	7.777	0.545	−5.631	67.809
0.05	2	7.165	0.196	−2.029	72.043	7.817	0.664	−3.341	68.979
0.05	3	7.165	0.196	−2.029	72.043	7.817	0.545	−3.048	67.823

The proposed framework are were envaluated in three aspects under different λ_1 and λ_2 in **Table 3**. Five different values of λ_1 and two values of λ_2 wereconsidered. The lowest average duration after using framework is obtained by $\lambda_1 = 0.03$ and $\lambda_2 = 3$, which is 67.792, while the average velocity, average acceleration and average treatment effect is 7.791, 0.545 and -5.633. It is obvious that after using our framework, the duration of the lane change can be significantly decreased and the efficiency of the lane change can be enhanced under the premise of lane change safety.

In order to visiulize the performance of the framework, data of 50 vehicles were used to test set, λ_1 is set to be 0.03 while λ_2 is set to be 3. **Figure 8** is the optimization results of 50 lane change vehicles. The red dots are suggested velocity while the blue dots are real velocity. The triangle points with different color are corresponding TTC values.

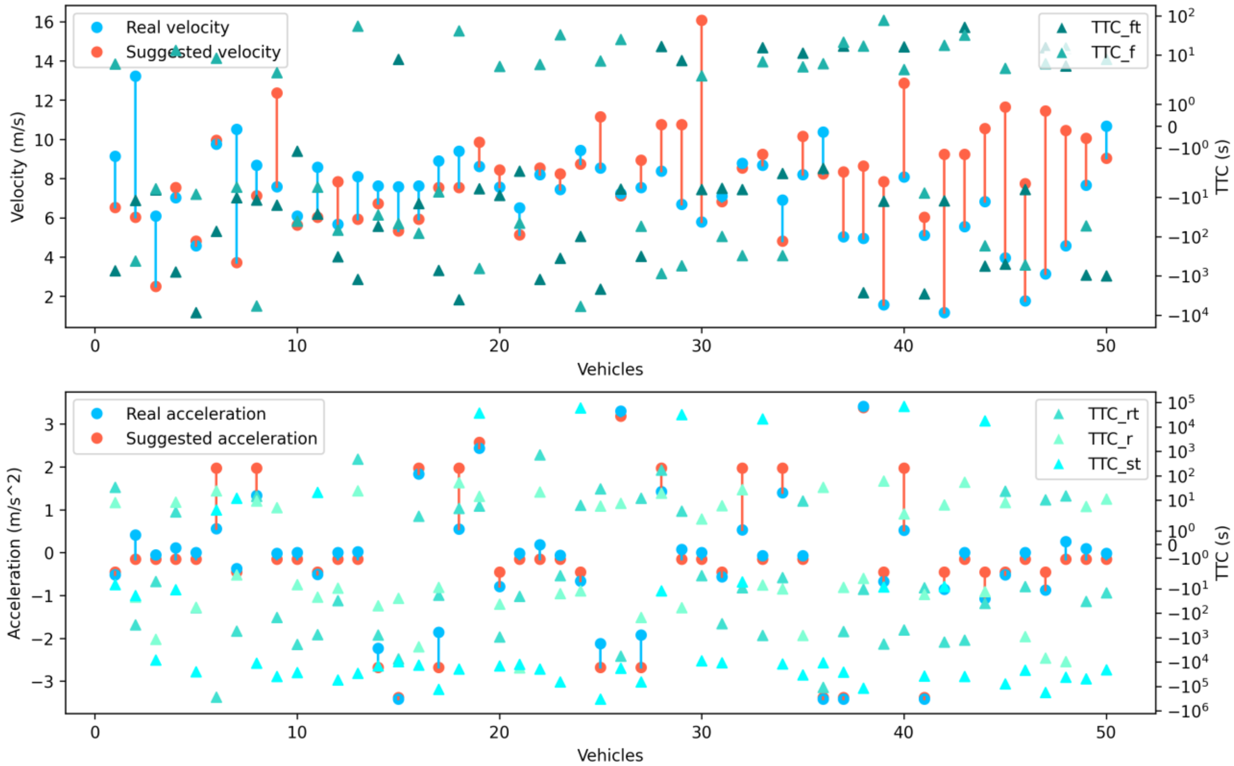

Figure 8 Change in velocity and acceleration before and after using framework

It can be noticed that the TTCs are all within the safe interval, which is 2.5 seconds. For vehicles with low velocity, suggested velocity tends to be higher than the real velocity. The acceleration after using the framework is similar to the acceleration without using the framework.

CONCLUSIONS

This study focus on the discretionary lane-changing process. A causal inference based speed control framework was proposed to help improve the lane change behavior from the aspect of efficiency, safety and eco-friendly, which can be denoted as velocity, TTC and acceleration respectively. Causal inference is applied to quantify the causality of variables. The causal effect of each variable on the lane change duration was calculated by using linear model and a placebo refutation test was performed. Conditional average treatment effect of treatment variable under covariates were also calculated by using causal forest. Two optimization problems were raised and solved in the framework to find optimal values of treatment variable and covariates. The final validation results show that the framework can improve the efficiency and environmental friendliness of lane change while ensuring the safety of lane change.

Along the stream of this study, several elements of future research can be identified. (1) More lane change parameters can also be added to the algorithm which can be treated as covariates, such as lane change angle or interaction variables of the target vehicle with surrounding vehicles. (2) A more generl framework can also be applied in Cooperative Vehicle Infrastructure System (CVIS). The framework proposed in this paper can be applied to the lane change system optimization module in collective intelligent decision making and cooperative control of traffic groups to improve the performance of lane change decisions.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Yi Zhao, Zhen Zhou; data collection: Zhen Zhou, Minghao Li, Yuyang Bao; analysis and interpretation of results: Yi Zhao, Zhen Zhou; draft manuscript preparation: Zhen Zhou, Minghao Li, Yuyang Bao. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

[1] Yang M, Wang X, Quddus M. Examining lane change gap acceptance, duration and impact using naturalistic driving data[J]. Transportation Research Part C: Emerging Technologies. 2019, 104: 317-331.

[2] Ahmed I, Xu D, Rouphail N, et al. Lane Change Rates at Freeway Weaving Sites: Trends in HCM6 and from NGSIM Trajectories[J]. Transportation Research Record: Journal of the Transportation Research Board. 2019, 2673(5): 627-636.

[3] Arbis D, Dixit V V. Game theoretic model for lane changing: Incorporating conflict risks[J]. Accident Analysis & Prevention. 2019, 125: 158-164.

[4] Zhang X, Sun J, Qi X, et al. Simultaneous modeling of car-following and lane-changing behaviors using deep learning[J]. Transportation Research Part C: Emerging Technologies. 2019, 104: 287-304.

[5] Li M, Li Z, Xu C, et al. Short-term prediction of safety and operation impacts of lane changes in oscillations with empirical vehicle trajectories[J]. Accident Analysis & Prevention. 2020, 135: 105345.

[6] Ali Y, Zheng Z, Mazharul Haque M, et al. Understanding the discretionary lane-changing behaviour in the connected environment[J]. Accident Analysis & Prevention. 2020, 137: 105463.

[7] Ali Y, Zheng Z, Haque M M, et al. A game theory-based approach for modelling mandatory lane-changing behaviour in a connected environment[J]. Transportation research. Part C, Emerging technologies. 2019, 106: 220-242.

[8] Ali Y, Haque M M, Zheng Z, et al. A hazard-based duration model to quantify the impact of connected driving environment on safety during mandatory lane-changing[J]. Transportation Research Part C: Emerging Technologies. 2019, 106: 113-131.

[9] Lin D, Li L, Jabari S E. Pay to change lanes: A cooperative lane-changing strategy for connected/automated driving[J]. Transportation research. Part C, Emerging technologies. 2019, 105: 550-564.

[10] Talebpour A, Mahmassani H S, Hamdar S H. Modeling Lane-Changing Behavior in a Connected Environment: A Game Theory Approach[J]. Transportation Research Procedia. 2015, 7: 420-440.

[11] Lyu N, Deng C, Xie L, et al. A field operational test in China: Exploring the effect of an advanced driver assistance system on driving performance and braking behavior[J]. Transportation Research Part F: Traffic Psychology and Behaviour. 2019, 65: 730-747.

[12] Cao P, Hu Y, Miwa T, et al. An optimal mandatory lane change decision model for autonomous vehicles in urban arterials[J]. Journal of Intelligent Transportation Systems. 2017, 21(4): 271-284.

[13] Mai T, Jiang R, Chung E. A Cooperative Intelligent Transport Systems (C-ITS)-based lane-changing advisory for weaving sections[J]. JOURNAL OF ADVANCED TRANSPORTATION. 2016, 50(5): 752-768.

[14] Luo Y, Xiang Y, Cao K, et al. A dynamic automated lane change maneuver based on vehicle-to-vehicle communication[J]. Transportation Research Part C: Emerging Technologies. 2016, 62: 87-102.

[15] Desiraju D, Chantem T, Heaslip K. Minimizing the Disruption of Traffic Flow of Automated Vehicles During Lane Changes[J]. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. 2015, 16(3): 1249-1258.

[16] Neuberg, Leland Gerson. Causality: Models, Reasoning, and Inference, by Judea Pearl, Cambridge University Press. Econometric Theory. 2003, vol. 19, no. 04.

[17] Ding, P. A paradox from randomization-based causal inference. Statistical Science. 2017, 32(3).

[18] Holland, PW. Statistics and Causal Inference, Journal of the American Statistical Association. 1986, 81:396, 945-960.

[19] Pearl, Judea. Causal diagrams for empirical research. Biometrika. 1995, 82(4), 702–710.

[20] Hitchcock, C., & Pearl, J. Causality: Models, reasoning and inference. The Philosophical Review. 2001, 110(4), 639.

[21] Li, C., Chan, S. H., & Chen, Y.-T. Who make drivers stop? towards driver-centric risk assessment: Risk object identification via causal inference. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020.

[22] You, T., & Han, B. Traffic accident benchmark for causality recognition. Computer Vision – ECCV. 2020, 540–556.

[23] Cao, D., Wu, J., Dong, X., Sun, H., Qu, X., & Yang, Z. Quantification of the impact of traffic incidents on speed reduction: A causal inference based approach. Accident Analysis & Prevention. 2021, 157, 106163.

[24] Ali, Q., Yaseen, M. R., & Khan, M. T. The causality of road traffic fatalities with its determinants in Upper Middle Income Countries: A continent-wide comparison. Transportation Research Part A: Policy and Practice. 2019, 119, 301–312.

[25] Yang, S., Ning, L., Cai, X., & Liu, M. Dynamic spatiotemporal causality analysis for network traffic flow based on transfer entropy and sliding window approach. Journal of Advanced Transportation. 2021, 1–17.

[26] Molavipour, S., Bassi, G., Čičić, M., Skoglund, M., & Johansson, K. H. Causality graph of vehicular traffic flow. arXiv preprint. 2020.
[27] Marchant J. Placebos: honest fakery. Nature. 2016 Jul;535(7611): S14-5.
[28] Athey S, Tibshirani J, Wager S. Generalized random forests. The Annals of Statistics. 2019 Apr;47(2):1148-78.
[29] Battocchi K, Dillon E, Hei M, Lewis G, Oka P, Oprescu M, Syrgkanis V. EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation. GitHub. 2019.
[30] Dong C, Zhang Y, Dolan JM. Lane-change social behavior generator for autonomous driving car by non-parametric regression in reproducing kernel hilbert space. In2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2017 Sep 24 (pp. 4489-4494). IEEE.
[31] Alexiadis V, Colyar J, Halkias J, Hranac R, McHale G. The next generation simulation program. Institute of Transportation Engineers. ITE Journal. 2004 Aug 1;74(8):22.
[32] Sharma A, Kiciman E. DoWhy: An end-to-end library for causal inference. arXiv preprint arXiv:2011.04216. 2020 Nov 9.