Evaluating the Impact of Heavy Vehicles on Lane-changing Decisions of Car Drivers: A Neural-network-based Methodology

Jian ZHENG[a], Koji SUZUKI[b], Motohiro FUJITA[c]

[a,b,c] *Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi, 466-8555, Japan*
[a] *E-mail: zhengjian1112@gmail.com*
[b] *E-mail: suzuki.koji@nitech.ac.jp*
[c] *E-mail: fujita.motohiro@nitech.ac.jp*

Abstract: From common driving experiences, car drivers are unwilling to follow heavy vehicles owing to speed and visibility obstructions, and most of them choose to change lanes to mitigate such obstructions. In literature, however, the impact of heavy vehicles on lane-changing decisions of following car drivers has rarely been investigated. In this study, a methodology for shedding light on such impact is proposed, which is based on two neural network models－the lane-changing model and the vehicle conversion model. Estimated by the large-scale trajectory data, the models both acquire high accuracy. By converting the nearest lead vehicles to heavy vehicles and detecting driving decisions of the following car drivers, the percentage of changed driving decisions under traffic conditions with the different proportion of heavy vehicles is clearly given, and the number of increased lane changes caused by heavy vehicles is quantitatively provided, as well.

Keywords: Artificial neural network, Lane-changing model, Vehicle conversion model, Heavy vehicle's impact, Car driver, Lane-changing decision

1. INTRODUCTION

Despite accounting for only a small proportion of all vehicular traffic, heavy vehicles have significant impact on traffic flow and traffic safety (Al-Kaisy et al. 2005, Moridpour et al. 2010, Stuster 1999). Besides, as reported in (Aghabayk et al. 2011, Peeta et al. 2005), the number of heavy vehicles has markedly increased over the past few decades and the trend is likely to continue at least over the next decade, in America and Australia. Therefore, more and more researchers have focused on the driving behavior of heavy vehicle drivers (Aghabayk et al. 2011, Moridpour et al. 2012, Moridpour et al. 2009, Toledo and Zohar 2007).

Owing to physical and operational characteristics, heavy vehicles impose physical and psychological effects on surrounding vehicles and drivers, especially for car drivers that directly follow them. From common driving experiences, it is also known that most car drivers are unwilling to follow heavy vehicles during their trips due to speed and visibility obstructions. In addition, under the heavy vehicle's influence car divers that immediately follow heavy vehicles tend to carry out different driving maneuvers even unsafe driving acts compared to those following cars, which are key causal factors for car-truck crashes (Stuster 1999). To mitigate the impact resulting from heavy vehicles, typically, most car drivers either maintain relatively great gaps or change lanes, which more or less lead to the capacity drop and additional driving risks. Furthermore, due to the fact that under congestion each lane change of heavy vehicles or their following vehicles may create new following vehicles, the number of cars that directly follow heavy vehicles is far more than the number of heavy vehicles. Therefore, with the increment of the number of heavy vehicles in vehicular traffic, much attention should be paid to the impact of heavy vehicle on surrounding vehicles, particularly on vehicles immediately following them.

Yoo and Green (1999) investigated the driver's behavior when following cars, trucks, and buses by using a driving simulator. Their findings show some differences in driving behavior when following different types of vehicles. However, credibility of their findings may be questionable, since they are based on a driving simulator. Besides, the differences in lane-changing behavior are not included in their study, as they restricted the passing behavior.

So far, to the best of our knowledge, there have been almost no studies relating to the impact of heavy vehicles on lane-changing decisions of car drivers. Conducting such studies may be restrained by two factors. First, to evaluate such impact, large-scale trajectory data are required to analyze the behavior of car drivers that immediately follow heavy vehicles. Second, appropriate methods for evaluating such impact have not been proposed. Although the microscopic traffic simulation can be used to investigate the impact of heavy vehicles on traffic flow (Moridpour et al. 2012), it is incapable of analyzing the impact on lane-changing decisions. The in-vehicle survey (Sun and Elefteriadou 2012) is feasible, but the cost for a large sample size is hardly affordable.

The objective of this paper is to attempt to quantitatively evaluate the impact of heavy vehicles on lane-changing decisions of car drivers. The methodology proposed in this study is based on two models－a neural network model of lane-changing decisions and a neural network model of vehicle conversion. The lane-changing model can completely take effects of surrounding vehicles in the current and adjacent lanes into consideration, and the vehicle conversion model is able to realistically convert movements of cars or motorcycles to those of heavy vehicles. By converting the nearest lead vehicles to heavy vehicles and detecting driving decisions of the following car drivers, the percentage of changed decisions under traffic conditions with the different proportion of heavy vehicles is clearly given, and the number of increased lane changes caused by heavy vehicles is quantitatively provided, as well. Additionally, in this study we also discuss some issues associated with the use of neural networks, such as the number of neurons in hidden layers and the sampling interval from the continuous vehicle trajectory, which may offer some guidelines to neural network users.

The paper is composed of six sections. Datasets used in this study are introduced in the subsequent section, followed by the description of the proposed methodology. Model estimation and evaluation results are presented in Sections 4 and 5, respectively. The last section is devoted to conclusions and further discussions.

2. DATASETS

The data used in this study were collected on a segment of U.S. Highway 101 in Los Angeles, California, by using eight video cameras that were mounted on a multi-story building. The study site is 640-meter long and covers an on-ramp and an off-ramp, as illustrated in Figure 1. The datasets are provided by Cambridge Systematic Incorporation for Federal Highway Administration, as a part of the Next Generation Simulation (NGSIM) program. Detailed information about observed vehicles (vehicle type and size, lane ID, two-dimension position, speed and acceleration) was extracted from video data, together with information about preceding and following vehicles.

Data reflecting congested traffic conditions in morning peak periods were collected between 7:50 am and 8:35 am on June 15, 2005. Traffic composition and traffic flow

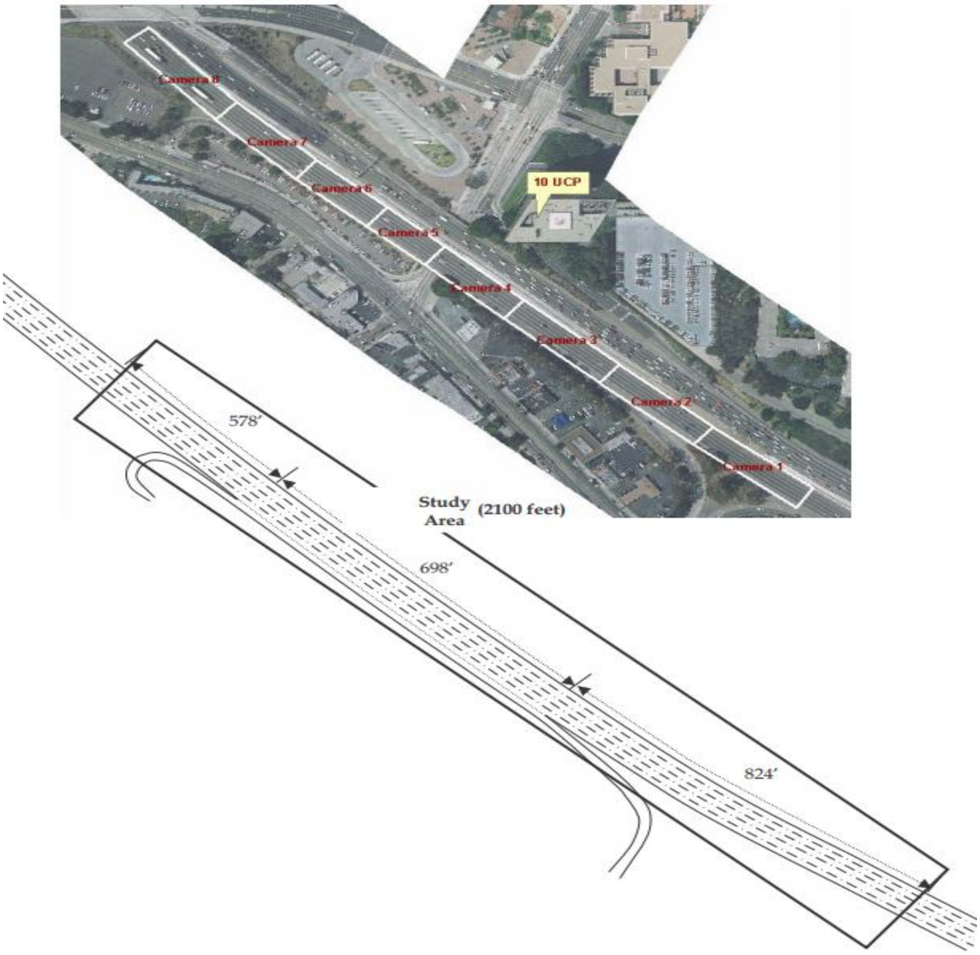

Figure 1 Illustration of the data collection site

Table 1 Traffic composition and traffic flow characteristics

Number of automobiles	Number of heavy vehicles	Number of motorcycles
5919 (97.0%)	137 (2.2%)	45 (0.8%)
Flow (veh/hr)	Speed (km/hr)	Density (veh/km)
8085	35.0	231
Level of service: E		

characteristics are listed in Table 1. Detailed data analysis and the data processing methodology are provided in NGSIM U.S. 101 Data Analysis Report (Cambridge Systematics Inc. 2005). Besides, in order to alleviate the noise in the data (Punzo et al. 2011), the moving-average filter for a duration of one second is applied to all vehicle trajectories before data analysis in this study.

3. METHODOLOGY

As we mentioned in the Introduction, appropriate methods for the evaluation of the heavy vehicle's impact on lane-changing behavior of car drivers have not been proposed. The original methodology developed in this study is to figure out responses of car drivers when the nearest lead vehicles in the same lane are adjusted to heavy vehicles and other surrounding traffic conditions keep the same. The methodology is based on two neural network models－

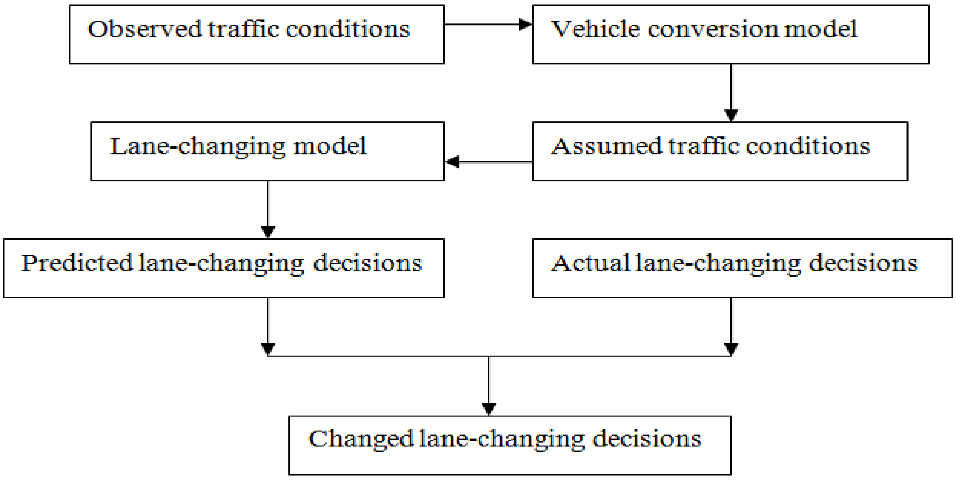

Figure 2 Methodology for the evaluation of the heavy vehicle's impact

the lane-changing model and the vehicle conversion model. As shown in Figure 2, first, according to the predetermined proportion of heavy vehicles, the vehicle conversion model is applied to convert the nearest lead cars or motorcycles to heavy vehicles. Then, under the assumed traffic conditions, the lane-changing model is employed to predict lane-changing decisions of following car drivers. Compared with actual lane-changing decisions, we can figure out how many drivers change their driving decisions under the impact of heavy vehicles. In this study, the lane-changing model is the most important component for the proposed methodology, and it is more complicated than the vehicle conversion model. For convenience, we introduced the lane-changing model first, although it is a little inconsistent with the contents in Figure 2.

3.1 Subject Vehicle and Its Surrounding Vehicles

In order to guarantee accuracy of the lane-changing and the vehicle conversion models, in this study we try to consider all possible factors that influence driving behavior. The subject vehicle and its surrounding vehicles are defined in Figure 3. It is noted that not only are the nearest lead and lag vehicles in the current and adjacent lanes taken into consideration but also the nearer lead and lag vehicles. To facilitate model specification, following variables are employed:

L_{id} = the ID of the current lane for the subject vehicle,
V = the instantaneous speed of the subject vehicle,
$T_{lead}^{nearest}$ = the type of the nearest lead vehicle,
$RV_{lead}^{nearest}$ = the relative speed between the nearest lead vehicle and the subject vehicle,
$SG_{lead}^{nearest}$ = the space gap between the nearest lead vehicle and the subject vehicle,
T_{lead}^{nearer} = the type of the nearer lead vehicle,
RV_{lead}^{nearer} = the relative speed between the nearer and the nearest lead vehicles,
SG_{lead}^{nearer} = the space gap between the nearer and nearest lead vehicles.

Accordingly, $T_{lag}^{nearest}, RV_{lag}^{nearest}, SG_{lag}^{nearest}, T_{lag}^{nearer}, RV_{lag}^{nearer}, SG_{lag}^{nearer}$ are variables with respect to the subject vehicle, the nearest and nearer lag vehicles.

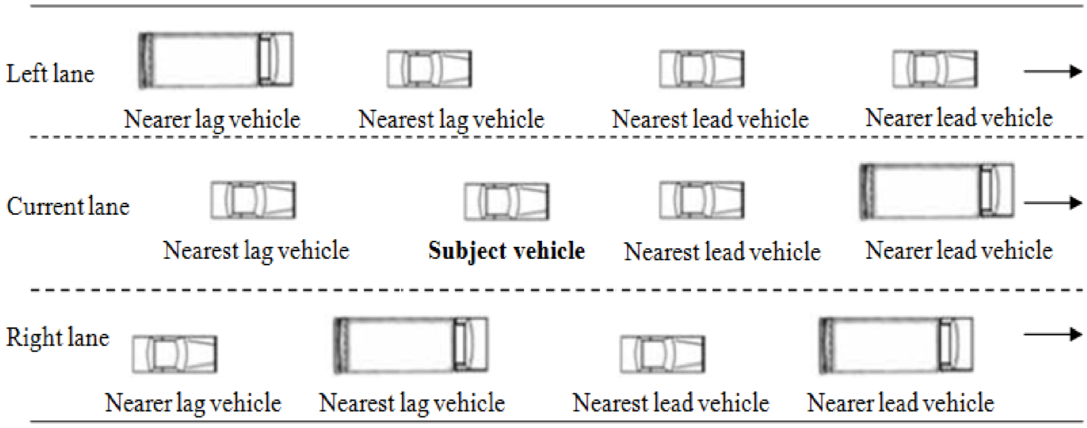

Figure 3 Illustration of the subject vehicle and its surrounding vehicles

3.2 Lane-changing Model

Since 1990s, there has been an increased interest in a wide variety of disciplines concerning the application of artificial neural networks (Adeli 2001, Hagan et al. 1996, Kalogirou 2000, Kalyoncuoglu and Tigdemir 2004, Karlaftis and Vlahogianni 2011, Rafiq et al. 2001). Typically, two advantages contribute to the popularity of NN models. One of them is that neural networks are able to handle noisy data and approximate any degree of complexity in nonlinear systems. Another advantage is that NN models do not require any simplifying assumptions or prior knowledge of problem solving, compared with statistical models. For example, in regression models, we have to specify the underlying relationship (linear, polynomial, exponential, rational, etc.) between independent and dependent variables before model estimation. However, such specifications are not necessary for inputs and outputs in neural network models.

Hunt and Lyons (1994) used two sorts of neural networks to attempt to model driver's lane-changing behavior. In the prediction-type network, drivers are unable to change lanes, due to that the model is trained by the simulated data from an individual subject driver where the portion of lane-changing instances is too small. The application of a classification-type neural network is considered to be viable. The model is able to correctly classify a very high proportion of examples during training for both simulated and observed data. However, it is also noted that misclassification of unseen driving examples is highly significant in the testing process. Besides, only the effect of the position of surrounding vehicles on the subject vehicles is considered in the classification-type network. The relative speed and the vehicle type of surrounding vehicles that also greatly influence lane changing are not taken into account (Moridpour et al. 2010).

The neural network model of lane-changing decisions developed in this study is a typical feed-forward neural network with five layers, as displayed in Figure 4. In the input layer, there are three input vectors and each one separately connects a hidden layer. Elements in each input vector are the variables associated with the vehicles in the left (L), current (C) and right (R) lane respectively, where

$$\begin{aligned} IP^L \text{ or } IP^R = \{ & T_{lead}^{nearest}, T_{lead}^{nearer}, RV_{lead}^{nearest}, RV_{lead}^{nearer}, SG_{lead}^{nearest}, SG_{lead}^{nearer}, \\ & T_{lag}^{nearest}, T_{lag}^{nearer}, RV_{lag}^{nearest}, RV_{lag}^{nearer}, SG_{lag}^{nearest}, SG_{lag}^{nearer}\}, \end{aligned} \quad (1)$$

$$IP^C = \{L_{id}, V, T_{lead}^{nearest}, T_{lead}^{nearer}, RV_{lead}^{nearest}, RV_{lead}^{nearer}, SD_{lead}^{nearest}, SD_{lead}^{nearer}, T_{lag}^{nearest}, RV_{lag}^{nearest}, SD_{lag}^{nearest}\}. \quad (2)$$

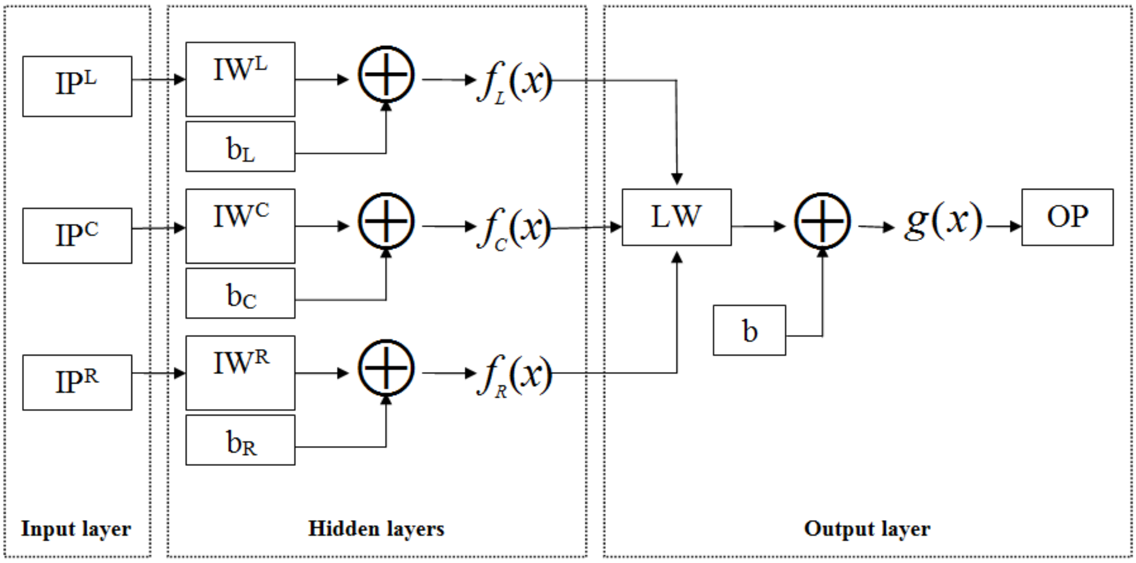

Figure 4 The neural network of lane-changing decisions

The transfer function in the hidden layers is the Hyperbolic tangent sigmoid function:

$$f(x_i) = \frac{2}{1+\exp(-2*x_i)} - 1,\ i \in [L, C, R]. \tag{3}$$

This function is inherently nonlinear and produces outputs with upper and lower bounds, which is considered suitable for complicated systems. The number of neurons in each hidden layer will be determined in the model estimation process in the next section.

Based on lane-changing directions, three neurons are employed in the output layer. The transfer function adopted in the output layer is defined as follows:

$$g(x_i) = \frac{\exp(x_i)}{\sum_{j\in[L,C,R]} \exp(x_j)},\ i \in [L, C, R]. \tag{4}$$

Such design provides the exact probability of making each decision, unlike the pattern recognition network where outputs are represented by zero or one.

The Levenberg-Marquardt back-propagation algorithm (Hagan et al. 1996) is used for training owing to its fast speed. The mean square error is taken as a measure of the performance of the neural network model,

$$MSE = \frac{1}{N}\sum_{i\in[L,C,R]} (OB_i - OP_i)^2, \tag{5}$$

where OB and OP are the observed lane-changing decisions and the predicted probability. N represents the number of total observations.

In addition, IW, LW and b_i stand for input and layer weights and bias which are needed to be estimated.

3.3. Vehicle Conversion Model

In order to realistically reflect the movements of heavy vehicles under assumed traffic conditions, a vehicle conversion model to convert the nearest lead vehicles to heavy vehicles is developed in this subsection. As the vehicle type variable is incorporated in the lane-changing model, the vehicle type of the nearest lead vehicles can be easily set to the heavy vehicle type. In what follows, we just need to adjust the speed of the nearest lead vehicles to the heavy vehicle speed.

The vehicle conversion model is also based on an artificial neural network, where three layers are used. One input vector is in the input layer and contains ten elements relating to the traffic conditions in the current lane — the lane ID of the subject vehicle, the vehicle type, the instantaneous speed and gap of its nearest and nearer lead vehicle and its nearest lag vehicle. The output vector is the speed of the heavy vehicle.

The transfer function used in the hidden layer is the same as Equation (1) and the one in the output layer is the log-sigmoid function,

$$g(x) = V_{\max} * \left(\frac{2}{1 + \exp(-x)} \right), \tag{6}$$

where V_{max} is the maximum speed among heavy vehicle samples.

The training method and the model performance measure are the same as in the lane-changing model.

4. MODEL ESTIMATION

Depending on the purpose, lane changing is categorized as being either mandatory or discretionary (Gipps 1986). Typically, mandatory lane changing is executed when the driver must leave the current lane to maintain the desired route. Discretionary lane changing refers to cases in which the driver changes lane to improve driving conditions, such as overtaking slow vehicles, passing heavy vehicles, avoiding traffic near an on-ramp. In addition, lane-changing maneuvers are different for different types of vehicles (heavy vehicle or car) (Aghabayk et al. 2011, Moridpour 2010).

In this study, we only discuss the discretionary lane changing of car drivers and a lane change is assumed to be instantaneous without duration. The instance in which a vehicle crosses over the current lane line in the next time step is treated as a lane-changing instance. Due to that lane-changing instances are greatly outnumbered by car-following or free-flow instances during a trip, it is impossible to train the neural network model by using continuous trajectories of some lane-changing vehicles. This is also the reason for the failure of the prediction-type network in (Hunt and Lyons 1994).

To mitigate this issue, in this research samples are drawn from continuous trajectory data. The sampling interval ΔT is defined in Figure 5, where a vehicle changes from lane 2 to lane 1. For lane-changing vehicles, the lane-changing instances are certainly included in the total samples and non-lane-changing instances before-and-after the lane-changing instances at each ΔT time interval are also selected. It is noted that with the small sampling interval the number of non-lane-changing samples will be increased. However, accuracy of lane-changing model for lane-changing samples is expected to be reduced.

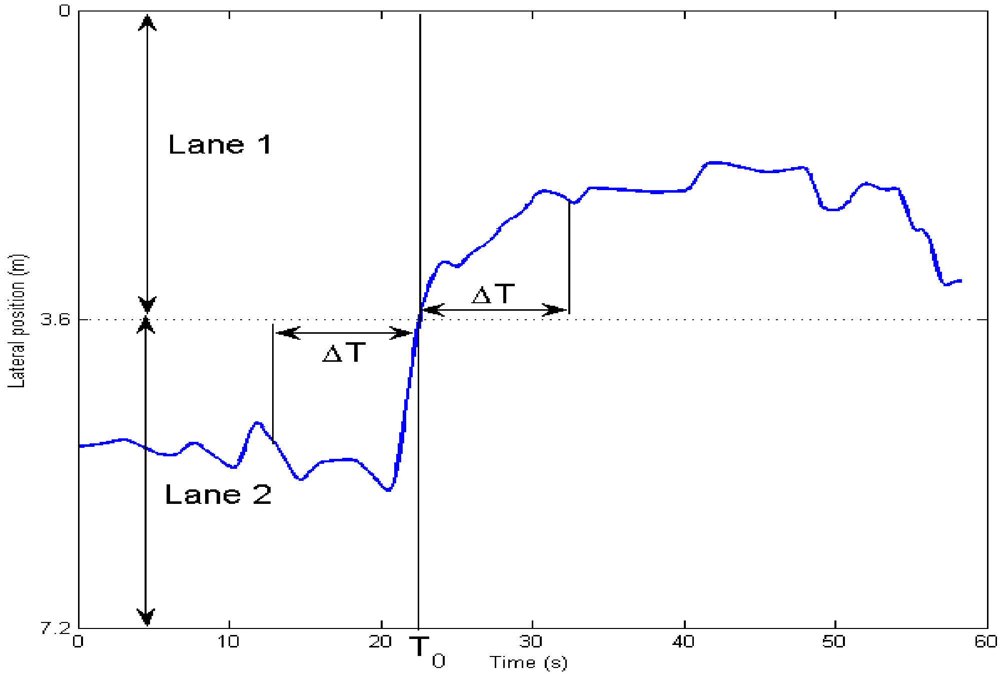

Figure 5 The definition of the sampling interval ΔT

In general, as the number of neurons used in the hidden layer increases, accuracy of the neural network model also increases in the model estimation process. However, the cost spent during model estimation and the over-fitting concern is also increased with the increment of the number of neurons in the hidden layer (Yin et al. 2003).

To make a better trade-off, by using the trajectory data collected between 7:50 and 8:05, the performance of lane-changing models with samples drawn at different time intervals and with the different number of neurons in the hidden layer is examined in the next subsection, where the result is the best one in ten simulation runs. Here, we mention that choosing the best result from several simulation runs is a typical way to avoid partial optimization caused by model initialization in the training process. In this study, one of the most widely used training methods in neural networks, the Levenberg-Marquardt back propagation algorithm, is adopted in all proposed models. In fact, checking the difference of MSE among 10 simulation runs is just to examine the consistency of the used training algorithm, which is beyond the topic of this research. Besides, the models are implemented in MATLAB 7.

4.1 Sensitivity Analysis of the Lane-changing Model

Estimated by samples drawn at different time intervals, Figures 6 and 7 present the performance of lane-changing models with different number of neurons in each hidden layer, where for an actual lane-changing sample if the predicted probability is more than 50% the prediction is considered to be correct.

From Figure 6, it shows that with the longer sampling interval, accuracy of the lane-changing model for lane-changing samples is increased. By using more neurons in each hidden layer, accuracy of the model for lane-changing samples is improved, as well. We also mention that, although by using the 15-second or 20-second sampling interval the percentage correctly predicted is higher than that by using the shorter sampling interval, the proportion of lane-changing samples to non-lane-changing samples is not consistent with the

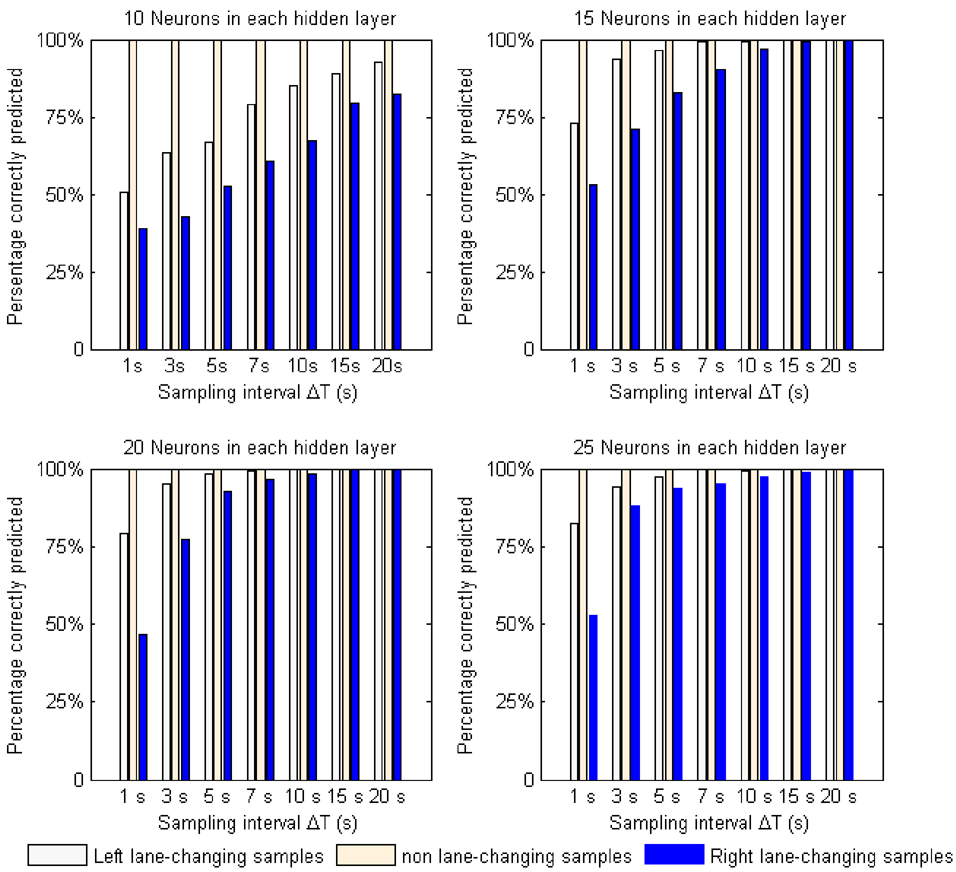

Figure 6 Accuracy of lane-changing models

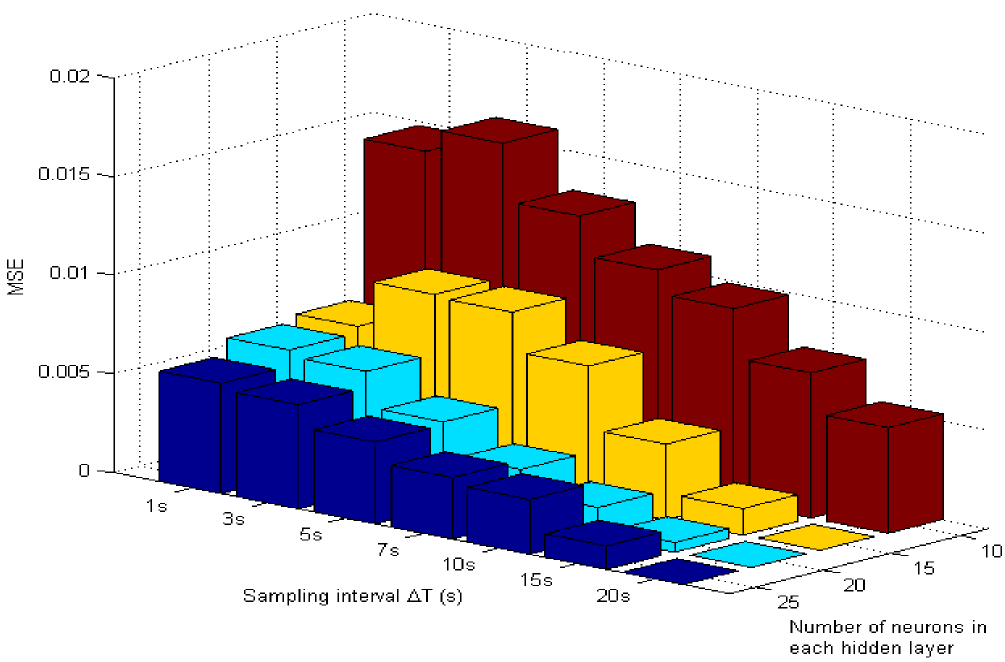

Figure 7 The mean square error of lane-changing models

real traffic where lane-changing instances are outnumbered by non-lane-changing instances. In addition, the models perform better for left lane-changing samples than for right lane-changing samples. Apart from the difference between the number of left and right lane-changing samples, the left and right lane-changing decisions are incentivized by different motivations, which also leads to the different prediction accuracy.

By the results in Figure 7, we can see that with the decrease in the number of total samples the mean square error of the models is also decreased. Moreover, the models with 20 neurons or 25 neurons in each hidden layer seems to be more stable than those with less neurons in each hidden layer. When the sampling interval is more than 7 seconds, the mean square error of the model with 20 neurons in each hidden layer is smaller than the model with 25 neurons in each hidden layer.

4.2 Estimation Results of the Lane-changing Model

Thoroughly considering the effects of the sampling interval and the number of neurons in each layer as shown in Figures 6 and 7, in the following analysis the sampling interval is taken as 10 seconds and 20 neurons are employed in each hidden layer of the lane-changing model. In all, 4618 samples are drawn from 858 lane-changing vehicles during 45-minute recording periods.

Table 2 The number of samples and the percentage correctly predicted

	Left lane-changing samples	Non-lane-changing samples	Right lane-changing samples	Total samples
Number of samples	801	3404	413	4618
Percentage correctly predicted	95.76%	99.69%	87.65%	98.52%

Accuracy of the lane-changing model for lane-changing, non-lane-changing samples and total samples are listed in Table 2. We can see that the percentage correctly predicted by the lane-changing model is very high in the model estimation process.

4.3 Estimation Results of the Vehicle Conversion Model

In order to represent realistic movements of heavy vehicles, the vehicle conversion model is trained by using heavy vehicle trajectory data. With the 10-second sampling interval, the number of samples drawn from trajectories of 121 heavy vehicles is 1189. The maximum speed, V_{max} in Equation (4), is set to 20 m/s. The number of neurons used in the hidden layer and the output layer is 10 and 1.

Figure 8 shows accuracy of the vehicle conversion model. With the correlation coefficient of 0.976 between the predicted and actual speed, the model can realistically reflect the movements of heavy vehicles. Besides, the dotted line in Figure 9 presents the actual speed profile of a car and the solid line shows the speed profile after conversion, that is, the speed of the assumed heavy vehicle. We note that, on the whole, the speed after conversion is lower than the actual speed and the changing tendency of actual and converted speed is consistent, which demonstrates reasonability of the vehicle conversion model.

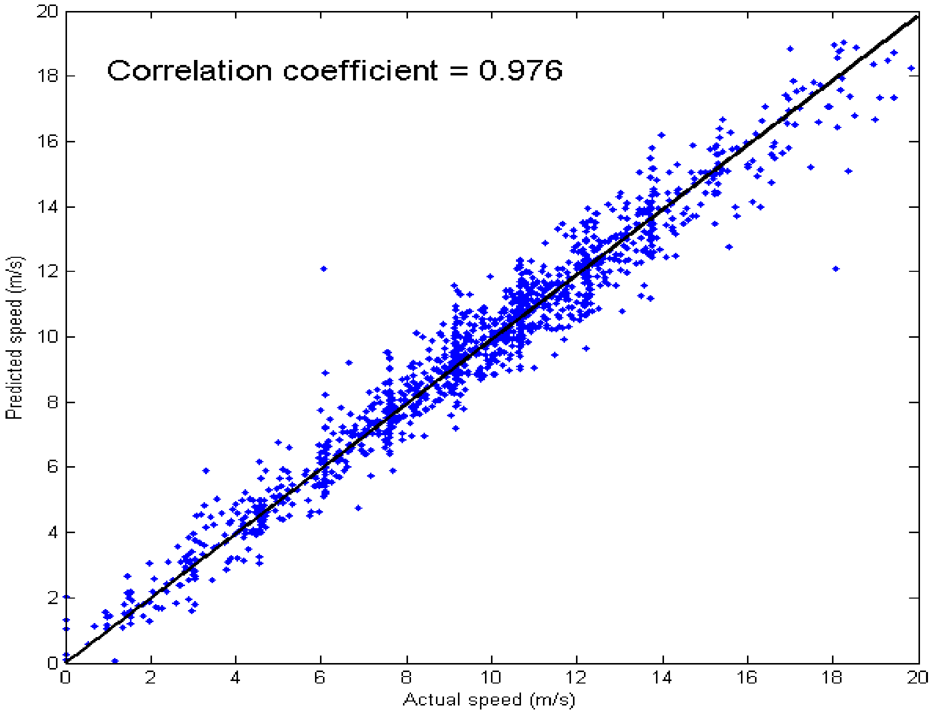

Figure 8 The scatter plot of the actual and predicted heavy vehicle speed

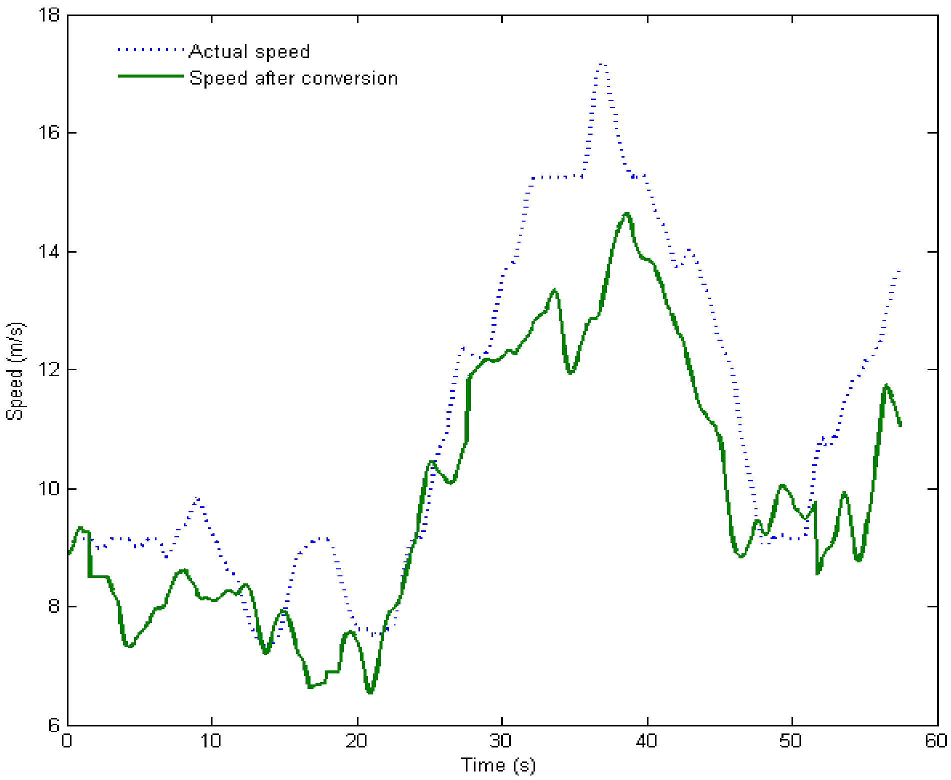

Figure 9 The actual speed of a car versus the speed after conversion

5. IMPACT EVALUATION

In the previous section, we can see that the lane-changing model and the vehicle conversion model both perform well in the estimation process. In fact, the performance of the models largely depends on the number of parameters adopted. Using too many parameters is prone to over-fitting results in model application (Hollander, Y. and Liu, R., 2008, Zheng et al. 2012a). In our opinion, if the model is used to predict unknown traffic conditions, we have to examine the over-fitting issue by using a new dataset. However, if we just want to quantitatively evaluate the effects of some independent variables on the dependent variable, namely sensitivity analysis of the model, it is not necessary to check the over-fitting issue. In this case, much attention should be paid to accuracy of the model. In fact, the proposed neural network lane-changing model is validated by using a new dataset and compared with a typical lane-changing model in one of our previous study (Zheng et al. 2013). Besides, a detailed analysis of the differences in right and left lane changes is also carried out.

In what follows, we use the proposed methodology to evaluate the impact of heavy vehicles on lane-changing decisions of car drivers. Typically, the proportion of heavy vehicles ranges from as low as 2% to as high as 25% of total traffic during the daytime (Al-Kaisy et al. 2002). As listed in Table 1, the proportion of heavy vehicles is 2.2% in the used data sets, and we set the proportion of heavy vehicles ranging from 2% to 26% in the following analysis.

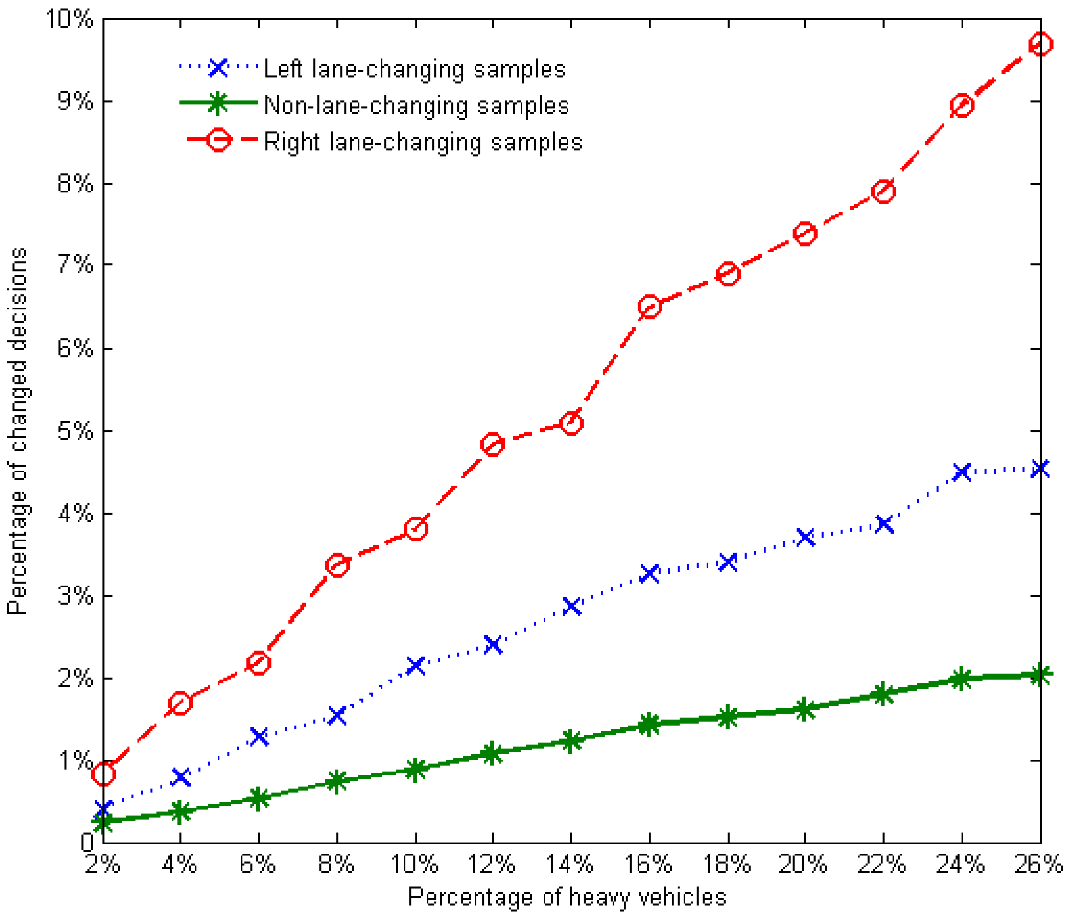

Figure 10 The percentage of changed decisions under the impact of heavy vehicles

First, we evaluate the impact of heavy vehicles on non-lane-changing decisions of subject car drivers. For non-lane-changing samples that are correctly predicted by the lane-changing model and have the nearest lead vehicles of cars or motorcycles in the current lane, we gradually raise the proportion of heavy vehicles from 2% to 26% by adapting the vehicle type of the nearest lead vehicles to be of the heavy vehicle type and converting the speed to the heavy vehicle speed based on the vehicle conversion model. As a result, the non-lane-changing decisions may be changed under the influence of vehicle conversion. And, the proportion of changed decisions can be detected by the lane-changing model.

In the same way, for left and right lane-changing samples that are correctly predicted by the lane-changing model and have the nearest lead vehicles of cars or motorcycles in the adjacent lane, the nearest lead vehicles in the left or right lane are also converted to heavy vehicles according to the predetermined heavy vehicle proportion. By such adaptation, some left or right lane-changing decisions may be rescinded. The lane-changing model can predict the driving decisions under such assumed traffic conditions.

The impact of heave vehicles on lane-changing and non-lane-changing decisions is displayed in Figure 10. From the figure, it is clear that, with the increment of the number of heavy vehicles in the current lane, among non-lane-changing samples the drivers deciding to change lanes gradually increase. When the proportion of heavy vehicles is 10%, around 1% non-lane-changing samples decide to change lanes. For left lane-changing samples, if 10% of the nearest lead vehicles in the left lane are heavy vehicles, more than 2% samples rescind their lane-changing decisions. Compared with non-lane-changing and left lane-changing samples, right lane-changing samples are more susceptible to the existence of heavy vehicles in the right lane. If the proportion of heavy vehicles in the right lane is 10%, nearly 4% samples decide to stay in the current lane. When the proportion of the preceding heavy vehicles in the corresponding lanes is 20%, the percentage of changed driving decisions among non-lane-changing samples, left lane-changing samples and right lane-changing samples is 1.6%, 3.7% and 7.4%, respectively.

By Figure 10, we note that for lane-changing samples heavy vehicles in the adjacent lanes can restrain the lane-changing behavior of car drivers. For non-lane-changing samples, however, heavy vehicles in the current lane force the following drivers to change lanes. Subsequently, we attempt to investigate the impact of heavy vehicles on the number of lane changes.

Because the samples used in the above analysis are drawn from lane-changing vehicles, in order to evaluate the impact of heavy vehicles on all vehicles, we have to extend the samples to non-lane-changing vehicles. By using the 10-second sampling interval, 18979 non-lane-changing samples are drawn from 4087 non-lane-changing vehicles. According to the percentage of changed decisions in Figure 10, the number of increased lane changes caused by heavy vehicles can be calculated as,

$$N = P_{non} \times N_{non} - P_{left} \times N_{left} - P_{right} \times N_{right} \quad (7)$$

where $P_{non}, P_{left}, P_{right}$ are the percentage of changed decisions among non-lane-changing samples and lane-changing samples and $N_{non}, N_{left}, N_{right}$ are the number of corresponding samples. The calculation results are displayed in Figure 11.

Figure 11 demonstrates that when the proportion of heavy vehicles is raised from 2% to 4%, the number of increased lane changes rises from 39 to 57, which is in agreement with the result in the observed datasets where the proportion of heavy vehicles is 2.2% and the number of lane changes caused by heavy vehicles is 46. Moreover, when the proportion of

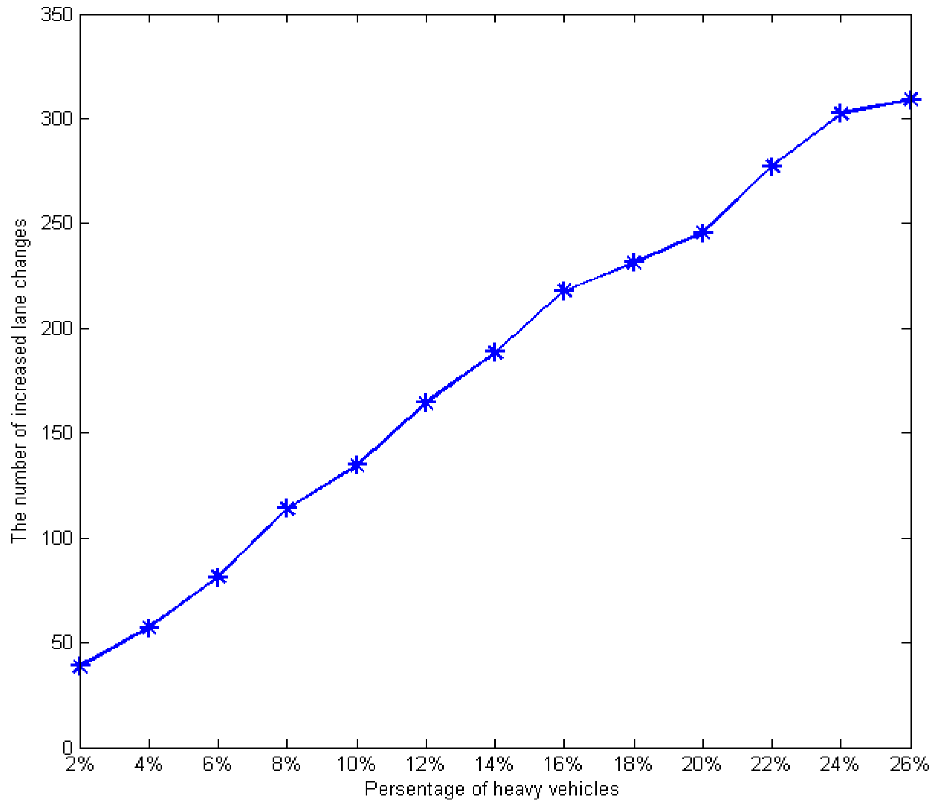

Figure 11 The number of increased lane changes under the impact of heavy vehicles

heavy vehicles is 10%, the number of increased lane changes is about 150, on the 610-meter long multilane freeway section during 45 minutes.

Here, we point out that extending the samples of lane-changing vehicles to the samples of total observed vehicles may lead to some inaccuracy in evaluation results. In fact, we have also attempted to estimate the lane-changing model by samples from all observed vehicles. However, the non-lane-changing samples significantly outnumber the lane-changing samples such that accuracy of the lane-changing model reduces dramatically.

6. CONCLUSIONS

According to common driving experiences, most car drivers usually carry out lane changes to mitigate speed and visibility obstructions from heavy vehicles, but the impact of heavy vehicles on lane-changing decisions has rarely been investigated in literature. In this study, a methodology for evaluating the impact of heavy vehicles on lane-changing decisions of car drivers is proposed. The methodology is based on two neural network models—the lane-changing model and the vehicle conversion model. Estimated by the large-scale trajectory data, the models both obtain high accuracy. Evaluation results show that: when the percentage of heavy vehicles in the corresponding lanes is 10%, around 1% non-lane-changing samples decide to change lanes; for the left lane-changing samples, more than 2% samples rescind their lane-changing behavior; the percentage of right lane-changing samples deciding to stay in the current lane is nearly 4%. Furthermore, when the proportion of heavy vehicles is raised from 2% to 4%, the number of increased lane changes rises from 39 to 57 on the 610-meter long multilane freeway section during 45 minutes. Such predicted result is consistent with the result in the observed datasets where the proportion of heavy vehicles is 2.2% and the number of lane changes caused by heavy vehicles is 46.

In this study, the impact of heavy vehicles on real traffic is evaluated from the point of view of following car drivers. It provides an essential supplement to the studies concerning heavy vehicles (Al-Kaisy et al. 2002, Aghabayk et al. 2011, Moridpour et al. 2009, 2010, 2012). Moreover, the findings in this research are helpful to the further understanding of driving behavior—car-following (Oguchi 2000, Zheng et al. 2012b) and lane-changing (Gipps 1986, Hunt and Lyons 1994). Besides, although the congested traffic conditions on a 610-meter long multilane freeway section are discussed in the current research, the proposed methodology can also be applied to uncongested traffic conditions and different types of roadways.

ACKNOWLEDGEMENT

The authors would like to thank the NGSIM program for data availability.

REFERENCES

Adeli, H. (2001) Neural networks in civil engineering: 1989 – 2000. *Computer-Aided Civil and Infrastructure Engineering* 16(2), 126 – 142.

Aghabayk, K., Moridpour, S., Young, W., Sarvi, M. and Wang, Y. (2011) Comparing heavy vehicle and passenger car lane-changing maneuvers on arterial roads and freeways. *Transportation Research Record: Journal of the Transportation Research Board* 2260, 94 – 101.

Al-Kaisy, A. F., Hall, F. L. and Reisman, E. S. (2002) Developing passenger car equivalents for heavy vehicles on freeways during queue discharge flow. *Transportation Research Part A: Policy and Practice* 36(8), 725 – 742.

Al-Kaisy, A., Jung, Y. and Rakha, H. (2005) Developing passenger car equivalency factors for heavy vehicles during congestion. *Journal of transportation engineering* 131, 514 – 523.

Cambridge Systematics Inc. (2005) NGSIM I-80 data analysis report. Technical report, Prepared for FHWA, U.S. Department of Transportation.

Gipps, P. G. (1986) A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological* 20(5), 403 – 414.

Hagan, M., Demuth, H., Beale, M. et al. (1996) *Neural network design*. Thomson Learning Stamford, CT.

Hollander, Y. and Liu, R. (2008) The principles of calibrating traffic microsimulation models. *Transportation* 35, 347–362.

Hunt, J. and Lyons, G. (1994) Modelling dual carriageway lane changing using neural networks. *Transportation Research Part C: Emerging Technologies* 2(4), 231 – 245.

Kalogirou, S. (2000) Applications of artificial neural-networks for energy systems. *Applied Energy* 67(1), 17 – 35.

Kalyoncuoglu, S. and Tigdemir, M. (2004) An alternative approach for modelling and simulation of traffic data: artificial neural networks. *Simulation Modelling Practice and Theory* 12(5), 351 – 362.

Karlaftis, M. and Vlahogianni, E. (2011) Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies* 19(3), 387 – 399.

Moridpour, S., Mazloumi, E., Sarvi, M. and Rose, G. (2012) Enhanced evaluation of heavy vehicle lane restriction strategies in microscopic traffic simulation. *Journal of Transportation Engineering* 138, 236 – 242.

Moridpour, S., Rose, G. and Sarvi, M. (2009) Modelling the heavy vehicle drivers lane changing decision under heavy traffic conditions. *Road and transport research* 18(4), 49 – 57.

Moridpour, S., Rose, G. and Sarvi, M. (2010) Effect of surrounding traffic characteristics on lane changing behavior. *Journal of Transportation Engineering* 136, 973–985.

Oguchi, T. (2000) Analysis of bottleneck phenomena at basic freeway segments: Car-following model and future exploration. *Journal of Infrastructure Planning and Management (JSCE)*, 49, 39–51. (in Japanese)

Peeta, S., Zhang, P. and Zhou, W. (2005) Behavior-based analysis of freeway car-truck interactions and related mitigation strategies. *Transportation Research Part B: Methodological* 39(5), 417 – 451.

Punzo, V., Borzacchiello, M. T. and Ciuffo, B. (2011) On the assessment of vehicle trajectory data accuracy and application to the next generation simulation (NGSIM) program data, *Transportation Research Part C: Emerging Technologies* 19(6), 1243 – 1262.

Rafiq, M., Bugmann, G. and Easterbrook, D. (2001) Neural network design for engineering applications. *Computers and Structures* 79(17), 1541–1552.

Stuster, J. W. (1999) The unsafe driving acts of motorists in the vicinity of large trucks, Technical report. Project report prepared for office of Motor Carrier Research.

Sun, D. J. and Elefteriadou, L. (2012) Lane-changing behavior on urban streets: An "In-vehicle" field experiment-based study, *Computer-Aided Civil and Infrastructure Engineering* 27, 525 – 542.

Toledo, T. and Zohar, D. (2007) Modeling duration of lane changes. *Transportation Research Record: Journal of the Transportation Research Board* 1999, 71 – 78.

Yin, C., Rosendahl, L. and Luo, Z. (2003) Methods to improve prediction performance of ANN models. *Simulation Modelling Practice and Theory* 11, 211 – 222.

Yoo, H. and Green, P. (1999) Driver behavior while following cars, trucks, and buses. Technical report, University of Michigan Transportation Research Institute.

Zheng, J., Suzuki, K. and Fujita, M. (2012*a*) Evaluation of car-following models using trajectory data from real traffic. *Procedia - Social and Behavioral Sciences* 43, 356 – 366.

Zheng, J., Suzuki, K. and Fujita, M. (2012b) Modeling acceleration and deceleration behavior in the HOV lane by using discrete choice theory. *Journal of Japan Society of Civil Engineers, Ser. D3* 68, 1163 – 1174.

Zheng, J., Suzuki, K. and Fujita, M. (2013) Predicting driver's lane-changing decisions using a neural network model. Under review.