## **Publish Information:**

Wei, D., and H. Liu (2012). "Analysis of Asymmetric Driving Behavior Using a Self-learning Approach." *Transportation Research Part B: Methodological* **46**(1), 1-14.

# Analysis of asymmetric driving behavior using a self-learning approach

Dali Wei<sup>1</sup>, Hongchao Liu<sup>2,3</sup>

<sup>1</sup>Department of Civil and Environmental Engineering, Texas Tech University, 10th and Akron, Lubbock, TX 79409 Email: <u>dali.wei@ttu.edu</u>; Fax: 1-(806)7424168; Telephone: 1-(806)7422801

<sup>2</sup> Department of Civil and Environmental Engineering, Texas Tech University, 10th and Akron, Lubbock, TX 79409

<sup>3</sup> School of Transportation Engineering, Hefei University of Technology, 193 Tunxi Rd, Hefei, China Email: <u>hongchao.liu@ttu.edu</u>; Fax: 1- (806)7424168; Telephone: 1-(806)7423523
\*Corresponding author

## ABSTRACT

This paper presents a self-learning Support Vector Regression (SVR) approach to investigate the asymmetric characteristic in car-following and its impacts on traffic flow evolution. At the microscopic level, we find that the intensity difference between acceleration and deceleration will lead to a 'neutral line', which separates the speed-space diagram into acceleration and deceleration dominant areas. This property is then used to discuss the characteristics and magnitudes of microscopic hysteresis in stop-and-go traffic. At the macroscopic level, according to the distribution of neutral lines for heterogeneous drivers, different congestion propagation patterns are reproduced and found to be consistent with Newell's car following theory. The connection between the asymmetric driving behavior and macroscopic hysteresis in the flow-density diagram is also analyzed and their magnitudes are shown to be positively related.

Keywords: Asymmetric Driving Behavior, Neutral Line, Hysteresis, Support Vector Regression

## 1. Introduction

The asymmetry in car-following has been well recognized, but few studies have been conducted to investigate its impact on traffic flow. The asymmetry stem from the difference of intensity (magnitude) between acceleration and deceleration is based upon the fact that the acceleration under positive relative speed of a given magnitude is not as intense as deceleration under a negative relative speed of the same magnitude (Leutzbach, 1988). This asymmetric characteristic is easy to understand since drivers pay closer attention to decreases than to increases in spacing. The intensity difference has been well observed and verified by Leutzbach (1988), Aron (1988) and Gong. (2008). Recently, Tordeux (2010) also verified the intensity difference between acceleration from Next Generation SIMulation (NGSIM) data.

Another well-known asymmetric characteristic is so-called hysteresis observed in stop-and-go traffic. Drivers are used to having a larger headway when accelerating compared to decelerating given the same speed in the hysteresis. Zhang (1999) proposed a theoretical approach to model the macroscopic hysteresis phenomenon observed from real traffic. He classified the traffic flow into three types: acceleration, deceleration, and equilibrium to obtain speed-concentration and occupancy relationships. Recently, Laval and Leclercq (2010) presented a measurement of hysteresis accounting for the non-stationary states and identified four types of hysteresis including strong, weak, negligible and negative. From the microscopic angle, Yeo (2008) presented an asymmetric driving behavior theory and explained some prominent macroscopic traffic phenomena from asymmetric properties. Ahn et al. (2012) also investigated the microhysteresis by comparing the equilibrium and observed spacing. The magnitude was found to be smaller in positive and larger in negative loops when measured with the equilibrium spacing.

In this research, the authors attempt to use a self-learning approach, namely Support Vector Regression (SVR), to further analyze the asymmetric property. SVR is superior at discovering the inherent relations among the variables in a dataset. Given a trajectory dataset, the asymmetric characteristics including both the intensity difference and hysteresis can be well tracked and learned by the SVR owing to its excellent learning ability.

Traditionally, the data analysis, aiming to provide an informational description of the data, usually adopts a multivariable regression approach, assuming a linear or nonlinear relation among variables. However, such a structural assumption on the mathematical formulation is usually appointed by experience or prior knowledge, which may lead to serious bias (Hair et al., 1998). If a linear form is adopted, the analytical model may be too simple to represent the underlying complex relations among the variables. But if applying a high-order form, the model might become excessively complex and over-fitted, which exaggerates minor fluctuations in the data and results in poor generalization performance due to the Empirical Risk Minimization (ERM) principle (Hastie et al., 2003).

The SVR approach overcomes the aforementioned drawbacks by applying the structural risk minimization (SRM) principle, which minimizes an upper bound on the expected risk and enables the tradeoff between the model complexity and empirical errors (Vapnik, 1995). This difference equips SVR with greater generalization abilities and allows a superior performance compared to conventional data analysis (Brown et al., 2000; Burbidge et al., 2001; Huang et al., 2007).

Based on the NGSIM data of U.S. Highway 101, both the intensity difference and hysteresis are analyzed from individual to heterogeneous drivers. At the micro-level we analyze the asymmetric property for an individual driver from three quantitative aspects: the neutral states of the intensity difference, the impacts of engaging vehicles' states on the hysteresis, and the relation between the intensity difference and the hysteresis. Then we extend the analysis to the macro-level with an emphasis on how the microscopic asymmetric behavior affects the congestion propagation patterns and the macroscopic hysteresis in the flow-density diagram.

The rest of the paper is organized as follows: the asymmetric car following model based on SVR is proposed in section 2. Section 3 presents the model's training and validation with the NGSIM data. A systematic analysis of the asymmetric property and discussion of the relations between the intensity difference and the hysteresis are given in Section 4. Section 5 extends the analysis to macro-level and analyzes how the micro-level driving behavior affects the congestion propagation patterns and the hysteresis in a flow-density diagram. Section 6 concludes this paper.

## 2. A self-learning approach for asymmetric driving behavior analysis

Support Vector Machine (SVM) is a universal learning method based on the statistical learning theory developed by Vapnik (1998) and Cherkassky (Cherkassky and Ma, 2004). SVM embodies the structural risk minimization (SRM) principle to minimize an upper bound on the expected risk. This feature equips SVM with a greater generalization performance from training samples to the unknown state space.

SVR is an application of SVM for regression problems. Given a set of training data  $(x_i, y_i), i = 1,2,3 \dots, l, x_i \in \mathbb{R}^d, y_i \in \mathbb{R}$ , a linear function regression function can be stated as:

$$f(x,\omega) = \omega^T \phi(x) + b \tag{1}$$

in a high-dimensional feature space  $\mathcal{F}$ , where  $\omega$  is a vector. Function  $\phi(x)$  maps the input x into a vector in the high-dimension space  $\mathcal{F}$ . The  $\mathcal{E}$ -intensive loss function proposed by Vapnik takes the form of:

$$L_{\xi}(y, f(x, \omega)) = \begin{cases} 0 & if |y - f(x, \omega)| < \varepsilon \\ |y - f(x, \omega)| - \varepsilon & otherwise \end{cases}$$
(2)

SVR performs a linear regression in space  $\mathcal{F}$  to minimize the empirical risk, as well as to reduce model complexity by minimizing  $\|\omega\|^2$ , which adds a capacity control and leads to a regularized risk function as:

$$R_{reg}(\omega) = \frac{1}{2} \|\omega\|^2 + R_{emp}(\omega) = \frac{1}{2} \|\omega\|^2 + \frac{1}{n} \sum_{i=1}^n L_{\mathcal{E}}(y_i, f(x, \omega))$$
(3)

Together with Eqs. (2) and (3), SVR has the following optimization problem in its dual form with a set of Lagrange Multipliers  $a_i$  and  $a_i^*$ :

$$\max L(\alpha_{i}, \alpha_{i}^{*}) = -\varepsilon \sum_{i=1}^{N} (\alpha_{i} + \alpha_{i}^{*}) + \sum_{i=1}^{N} (\alpha_{i}^{*} - \alpha_{i}) y_{i} - \frac{1}{2} \sum_{i,j=1}^{N} (\alpha_{i}^{*} - \alpha_{i}) (\alpha_{j}^{*} - \alpha_{j}) (\phi(x_{i}) \cdot \phi(y_{j}))$$

$$s. t. \begin{cases} y_{i} - f(x_{i}, \omega) \leq \mathcal{E} + \xi_{i} \\ f(x_{i}, \omega) - y_{i} \leq \mathcal{E} + \xi_{i} \\ \xi_{i}, \xi_{i}^{*} \geq 0, i = 1, ..., n \\ 0 \leq a_{i} \leq C, 0 \leq a_{i}^{*} \leq C \end{cases}$$
(4)

The maximization problem of Eq. (4) only utilizes the dot products of  $\phi(x_i) \cdot \phi(y_j)$ , and it suffices to know  $\phi(x_i) \cdot \phi(y_j)$  rather than as a mapping function of  $\phi(x_i)$ . With the kernel technique,  $\phi(x_i) \cdot \phi(y_j)$  can be transformed to a nonlinear kernel function  $K(x_i, x_j)$ . Eq. (4) can be further written as:

$$\max L(\alpha_{i}, \alpha_{i}^{*}) = -\varepsilon \sum_{i=1}^{N} (\alpha_{i} + \alpha_{i}^{*}) + \sum_{i=1}^{N} (\alpha_{i}^{*} - \alpha_{i}) y_{i} - \frac{1}{2} \sum_{i,j=1}^{N} (\alpha_{i}^{*} - \alpha_{i}) (\alpha_{j}^{*} - \alpha_{j}) K(x_{i}, x_{j})$$

$$s. t. \begin{cases} y_{i} - f(x_{i}, \omega) \leq \mathcal{E} + \xi_{i} \\ f(x_{i}, \omega) - y_{i} \leq \mathcal{E} + \xi_{i} \\ \xi_{i}, \xi_{i}^{*} \geq 0, i = 1, ..., n \\ 0 \leq a_{i} \leq C, 0 \leq a_{i}^{*} \leq C \end{cases}$$
(5)

At the optimal solution of Eq. (5), we obtain:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) \qquad s.t. \quad 0 \le \alpha_i \le C, 0 \le \alpha_i^* \le C$$
(6)

To note that, in Eq. (6) there is a coefficient for each of the training data. Those data with corresponding  $(a_i - a_i^*) \neq 0$  are called support vectors and contribute to the decision function. Given  $n_{sv}$  support vectors, Eq. (6) can be further written as:

$$f(x) = \sum_{i=1}^{n_{sv}} (\alpha_i - \alpha_i^*) K(x_i, x) \qquad s.t. \ 0 \le \alpha_i \le C, 0 \le \alpha_i^* \le C$$
(7)

Based on Eq. (7), the proposed SVR driving behavior model is shown in Fig. 1. The input of the model is a state vector  $x = (d(t), v_f(t), \Delta v(t))$  at time t, where the d(t) is the spacing between the two consecutive vehicles,  $\Delta v(t)$  is the relative speed of the two vehicles, and  $v_f(t)$  is the speed of the leading vehicle. The output of the model is the velocity of the following vehicle of  $v_f(t+T)$  at time t + T, where T is the correlation delay. Once the model is trained with samples,

an optimal estimation of the relation among multiple variables in car following behavior can be obtained as Eq. (8).

$$v_f(t+T) = \sum_{i=1}^{n_{sv}} (\alpha_i - \alpha_i^*) K((d_i(t), v_{f,i}(t), \Delta v_i(t)), (d(t), v_f(t), \Delta v(t)))$$
(8)



Fig. 1 SVR based car following behavior model

## 3. Model training and validation with NSGIM data

This section discusses the training and validation process of the SVR approach with NGSIM data. The data selection process is introduced first to choose the suitable data for tracking the driving behavior through traffic disturbances. Then validations including both model test and simulation test are performed to evaluate the accuracy. The asymmetric property of the model is also verified in this section.

## 3.1 NGSIM data preprocessing

A 15 min (7:55-8:05 AM) trajectory data from lane one of Highway 101 (FHWA, 2007) is used in our study. The vehicle pairs whose velocities were less than 1 m/s at some time during the studying period were selected to guarantee these trajectories can represent the driving behavior in typical stop-and-go conditions. A Total of 106 vehicle trajectory data are selected and marked as red dots in Fig. 2. A sample of selected vehicles' velocities and spacing are plotted in Fig. 3.

Each trajectory was evenly divided into two datasets. One is used for model training and the other for model validation including both model test and simulation test. Data samples in each trajectory with even time indexes are assigned as the training samples, while that with odd time indexes are validation samples. It ensures the training dataset covers sufficient information under different driving states reflected in the trajectory.





Fig. 3 A sample of selected leading and following vehicles' velocity and spacing

## 3.2 Model Training and Validation

The model is then trained with the calibration dataset for each trajectory using the LibSVM toolbox (Chang and Lin, 2011), where the regulation parameter, the coefficient of the kernel function and the loss function parameter are determined with the 5-fold cross-validation method (Cherkassky, 2004). The delay lag of *T* is set to be 1.0 second in this experiment, which is close to the reaction time in real driving behavior and adopted in a number of car following model calibration and validation studies (Bham and Benekohal, 2004; Chakroborty and Kikuchi, 1999; Johansson and Rummer, 1971).

To validate the model after training process, experiments including both model test and simulation test are performed (Sakda and Hussein, 2007). The model test is a static test, in which each individual sample in the validation set is fed into the model. The output is compared with the output of velocity in the validation dataset. Root Mean Square Error is used as one of the error indicators, which measures the deviation of the predicted value from the field data.

Root Mean Square Error = 
$$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(y_n^s - y_n^0)^2}$$
 (9)

where  $y_n^s$  is the simulated value at time n,  $y_n^0$  is the corresponding field data, and N is the number of samples. Both positive and negative errors regarding  $y_n^s - y_n^0$  are also calculated as they indicate whether the model is underestimating or overestimating compared with the observed values.

Another error indicator is the Theil's inequality coefficient, which is defined as:

$$U = \frac{\sqrt{\frac{1}{N} \sum_{n=1}^{N} (y_n^s - y_n^0)^2}}{\sqrt{\frac{1}{N} \sum_{n=1}^{N} (y_n^s)^2} + \sqrt{\frac{1}{N} \sum_{n=1}^{N} (y_n^0)^2}}$$
(10)

The numerator of U is the RMS error and the denominator is the sum of mean squares of observed and predicted values. Theil's inequality coefficient provides a measurement of RMS error in relative terms, which always falls between 0 and 1. It is easy to see that if U = 0, there is

a perfect fit with  $y_n^s = y_n^0$ . If U = 1, the performance of the model is no better than a naïve guess (Bham and Benekohal, 2004).

The errors of the model test are listed in Table 1 including the minimum, median, mean and maximum errors. All the errors in the model test are very small. The median RMS error is 1.7909 and the Theil's inequality coefficient is only 0.0422. Even for the largest deviation, it is only 2.6301 for RMS and 0.0736 for Theil's inequality coefficient. The median negative RMS is 1.9329, which is slightly higher than negative RMS. The negative mean percent error for speed test is a little higher than the positive mean percent error, which means that our model slightly tends to underestimate the speed. Overall, the model test errors are very small, which demonstrates that the predicted output velocity of our model closely match the field data.

		Table 1. Model test error							
		Root Mean Square Error	Positive Mean Square Error	Negative Mean Square Error	Theil's inequality coefficient				
Velocity	Minimum	0.3181	0.3789	0.0660	0.0053				
(Km/h)	Median	1.7909	1.7113	1.9329	0.0422				
	Mean	1.9343	1.8581	1.9774	0.0444				
	Maximum	2.6301	2.7535	2.0388	0.0736				

Simulation test is performed as a supplement to the model test, in which the leading vehicle's velocity profile, the initial spacing and initial following velocity are the input from the field data. The successive inputs of the model are automatically updated by the simulation itself. The output from the simulation is then compared with the field trajectory including the position, velocity and acceleration. The results are listed in Table 2. An example of simulated trajectories for the trajectory shown in Fig. 3, as well as the deviations, is plotted in Fig. 4.

	Root Mean Square Error				Theil's inequality coefficient			
	Minimum	Median	Mean	Maximum	Minimum	Median	Mean	Maximum
Velocity (km/h)	0.7796	2.7673	3.0415	6.6343	0.0130	0.0548	0.0601	0.1167
Trajectory (meter)	1.5132	2.7322	3.8669	8.0319	0.0181	0.0452	0.0598	0.1209
Acceleration (meter/s <sup>2</sup> )	0.5380	0.7121	0.7759	1.3711	0.3070	0.4335	0.4406	0.6171
Deceleration (meter/s <sup>2</sup> )	0.4350	0.6255	0.6456	1.0573	0.3648	0.4498	0.4837	0.6132

It can be seen that the simulation errors tend to be slightly higher than model test errors since the inputs of the simulation are automatically updated and deviate from the field inputs. The median RMS error and Theil's inequality coefficient for the trajectory are only 2.7322 meters and 0.0452, respectively. The mean velocity errors are also very small. For the median of RMS error, it is 2.7322 km/h. The acceleration and deceleration are relatively higher than velocity and trajectory errors in terms of Theil's inequality coefficient. The results also do not indicate obvious unbalanced predictions between deceleration and acceleration.



Fig. 4 (a) Following trajectories of simulation result and field data (b) Deviations from the trajectory

## **3.3** Verification of the asymmetric property

In this section, the asymmetric characteristics including intensity difference and hysteresis will be verified followed by a systematic analysis in section 4. The intensity difference is tested by changing the relative speed while keeping the input of following speed and spacing constant. Fig. 5 shows the output acceleration of the model when the following vehicle's speeds are 20 km/h and 30 km/h. The intensity difference can be clearly seen from the figures. For example, in the case the following speed is at 20 km/h and the spacing is 7.5 meters, the deceleration of the following vehicle is about  $2 \text{ m/s}^2$  when the relative speed is negative 10 km/h. In contrast, the acceleration is about  $0.3 \text{ m/s}^2$  if the relative speed is positive 10 km/h. In this case, the deceleration intensity is larger than acceleration. However, when the speed is 20 km/h and the spacing is 12 meters, the acceleration output of fixing the spacing while changing the following vehicle's speed and the relative speed. The same intensity difference is also clearly demonstrated between acceleration and deceleration.

There is a so-called 'neutral space' shown as the blue dotted line in Fig. 5(a). At the 'neutral space', the intensity of acceleration and deceleration will be symmetric. If the spacing is larger than the 'neutral space', the acceleration intensity is always stronger than deceleration, and vice versa. A similar 'neutral speed' can be found in Fig. 5(d) as well. The neutral state is an essential feature of the driving behavior, which can indicate and reflect different driving styles. A detailed

discussion and analysis of the neutral state will be given in following sections for better understanding of the asymmetric characteristic.



Fig. 5 Accelerations produced by our model for different relative speed and headways.



Fig. 6 Positive and negative loop produced by our model

The hysteresis loops between acceleration and deceleration phases are verified by simulating a typical stop-and-go traffic scenario. The initial speeds of both vehicles and spacing are 30 km/h and 12 meters. Then the leading vehicle will begin to slow down at a constant deceleration, stop for 10 seconds, and speed up again at a constant acceleration until its velocity returns to 30 km/h. Fig. 6 plots the hysteresis loops produced by the model with different trajectories. The hysteresis loop shown in Fig. 6(a) is a typical positive loop. The spacing in acceleration is always larger than deceleration. Meanwhile, Fig. 6(b) shows a negative hysteresis loop produced by the model trained with the trajectory data of vehicle 494. However, the negative hysteresis loops are reproduced for only 6 training trajectories out of a total of 106 tested training trajectories, which

is consistent with Laval's finding that 66% of platoons have positive direction hysteresis phenomena, 20% negligible, and 14% negative hysteresis (Laval, 2011). Comparing the negative hysteresis loop with the positive strong-level one, it seems that the negative hysteresis is not significant and can be classified with the weak level which can be explained by stochastic behavior and vehicle mix (Yeo and Skabardonis, 2009). Thus the following analysis and discussion will be focused on the positive direction hysteresis.

## 4. Asymmetric behavior analysis for an individual driver

#### 4.1 Neutral line in asymmetric driving behavior

The so-called 'neutral space' and 'neutral speed' in the relative speed-acceleration diagram have been found in Fig. 5. For every space, there is a corresponding 'neutral speed' and for every speed, there also exists a 'neutral space'.

Fig. 7 shows the neutral spaces for different following speeds as blue dots. We call the red line in the 'neutral line', which separates the speed-space diagram into two areas. The area below the neutral line is where deceleration is stronger than acceleration. That means in this area, the driver tends to have a more intensive deceleration than acceleration given the same speed and space but opposite relative speed. We call this area the deceleration dominant area. The other area above the line is the acceleration dominant area where acceleration is more intensive than the deceleration.



We can also consider the speed on the neutral line as the desired speed for the corresponding space since the driver will neither accelerate nor decelerate at each point of the neutral line. If the leading vehicle keeps the velocity constant, the following vehicle will remain in this neutral state. From this perspective, the equilibrium relationship between the spacing and velocity in Newell's car following (NCF) model (Newell, 2002) can be explained and determined by the neutral line. Figure 8 further shows an example of the distribution of differences between acceleration and deceleration when the relative speeds are set to be 10 and -10 km/h. It can be seen that the neutral line (red curve) separates the plane into two regions. The farther from the line, the larger the difference grows. In the tested scenario, the largest difference in the acceleration dominance area is around 10 m/s<sup>2</sup> and -12 m/s<sup>2</sup> for the deceleration dominance area. The neutral line can be regarded as a basic measurement of the asymmetric property in driving behavior. In the

following sections, we will further show that the neutral line is directly associated with the hysteresis and the congestion propagation patterns in macroscopic traffic.



Fig 8. Acceleration and deceleration differences

#### 4.2 Hysteresis analysis in the space-speed diagram

This section analyzes the microscopic hysteresis phenomenon with its focus on how the states of the engaging vehicles affect the magnitude of the hysteresis. The impact of the leading vehicle's behavior on the asymmetry is investigated at first. The leading vehicle is simulated with different accelerations in order to reproduce hysteresis loops with different magnitudes. The simulation scenario is the same as described in section 3.3 except for the leading vehicle's accelerations. The results are shown in Fig. 9.



Fig. 9 Hysteresis loops with different decelerations and accelerations

It can be seen from the figure that the magnitude of the hysteresis loop is affected by the acceleration and deceleration of the leading vehicles. The internal space of the hysteresis becomes wider with the increase of the leading vehicles' acceleration and deceleration. In the example shown in Fig. 9, for the length of the internal space, the largest value is around 12

meters when the leading vehicle's deceleration is 4.5  $m/s^2$  and 8 meters for acceleration at 4.5  $m/s^2$ .

In addition to the magnitude of decelerations, the speed at which the leading vehicle begins decelerating also affects the hysteresis. Figure 10 shows the hysteresis loops reproduced when the leading vehicle begins decelerating at different initial speeds. We can see that the external shape is determined by the leading vehicle's maximum speed in stop-and-go traffic. The external length of the hysteresis on the speed axis is almost equal to the maximum leading speed. And the external length on the space axis grows with the increase of the speed length.



Fig. 10 Hysteresis Loops with different leading speeds

In the above discussion, the impact of the leading vehicle's behavior (acceleration, deceleration and speed) on the hysteresis phenomenon is analyzed. Next attentions will be directed to the following vehicle's impact on the hysteresis.

The effect of different starting speeds of the following vehicle is examined first. In the simulated scenario, the initial speed of the leading vehicle and spacing are 30 km/h and 12 meters. Then the leading vehicle begins to slow down at a constant deceleration, stops for 10 seconds, and speeds up again at a constant acceleration until its speed returns to 30 km/h. Fig. 11(a) shows the hysteresis with different starting speeds. We can see that no matter what the starting speeds are, the following vehicles will eventually drop into the hysteresis loop. And its shape is affected by the leading vehicles' states. For example, when the starting speed is 50 km/h and starting space is 15 meters, the following vehicles will decelerate then fall into the curve on the loop. For the starting speed of 10 km/h, although the leading vehicle is decelerating, the 10 km/h speed is safe for the following vehicle, it will accelerate and then keep up with the loop. The impact of the starting space of the following vehicle is also shown in Fig. 11 (b). Similarly, with different starting spaces, the following vehicles will eventually drop into the hysteresis loop.



#### 4.3 Neutral line in hysteresis

As two important indicators of the asymmetric driving behavior, the intensity difference and hysteresis have been investigated and analyzed thus far. In this section, we made an attempt to find how these two phenomena are related in the driving task.



Fig. 12 Illustration of the hysteresis and the neutral line

We put the neutral line and the hysteresis loops in one speed-space diagram shown in Fig. 12. It can be seen that the turning points on the hysteresis loop are always on the neutral line. The neutral line determined by different magnitudes of acceleration and deceleration separates the hysteresis loop. The top half of the loop belongs to the acceleration dominance area, and the bottom half belongs to the deceleration dominance area. When the following speed equals to zero, the stationary space on the deceleration branch of the hysteresis is less than the neutral space, which indicates the following vehicle stops at a jam space and does not accelerate until the spacing is larger than the neutral space.

The intensity difference and hysteresis loops are essentially in accordance with each other. The magnitude distribution of the acceleration and deceleration intensity shown in Fig. 8 can also be explained in the hysteresis loop. Since the points are calculated in equal intervals (1 second in our simulation), the changes in speed between the points equal to the values of the accelerations or decelerations. For example, in Fig. 12(b),  $d_1$  is the deceleration value when the driver

decelerates from 10 km/h to 4 km/h.  $a_1$  is the acceleration value when the driver accelerates from 10 km/h to 15 km/h. The acceleration near the neutral line is smaller than those far from it. This is also consistent with the distribution shown in the Fig. 8, in which the difference between acceleration and deceleration is always smaller near the neutral line.

Essentially, both kinds of asymmetric properties are related. The leading vehicle's states affect the magnitude of the hysteresis while the neutral line determines the tilt angle and position of the hysteresis in the speed-space diagram. The neutral line reflects the inherent driving behavior, while the states of the leading vehicles play the role of driving stimuli.

## 5. Asymmetric behavior analysis from a macroscopic view

In previous sections, we analyzed the asymmetric characteristic based on the trajectory of individual vehicles. Next, we will extend our investigation to the macro-level to depict how the microscopic asymmetry is related to macroscopic traffic phenomena, such as congestion propagation and hysteresis in the flow-density diagram.

## 5.1 Neutral line distribution

The neutral line distribution based upon the tested 100 vehicles' trajectories (6 negative hysteresis loops are excluded) are shown in Fig. 13. Obviously, there is a rising trend in the neutral space from low to high speed. Most frequent neutral spacing is between 6-12 meters when speed is between 0-10 km/h. When the speed goes up to 30-40 km/h, the most frequent neutral spacing shifts to 12-24 meters. The neutral states in lower speeds are more centered compared to those at higher speeds. When the velocity is 0 km/h, the neutral spaces are centered within 6 to 12 meters. The neutral space's distribution for velocities between 30-40 km/h is smoother and more evenly distributed between 9-36 meters compared with 10-20 km/h.

Up and down boundaries are found in the neutral line distribution shown as red lines in Fig. 15(a). We call the driver along the upper boundary a cautious driver, and the driver along the lower boundary an aggressive driver. This is reasonable since an aggressive driver always tends to seek a smaller neutral space than a cautious driver. For a cautious driver, the deceleration dominant area is larger compared with that of an aggressive driver. The cautious driver also exhibits a larger stationary spacing which is about 10 meters, while for the aggressive driver, it is around 6 meters.



#### 5.2 Impact on congestion propagations

Drivers identified by different neutral lines clearly show different asymmetric properties and follow different acceleration and deceleration trajectories in hysteresis. Considering such differences from a macroscopic level, the impact of different asymmetry on the propagations of congestion is discussed in this section.

We designed two test scenarios with a platoon of vehicles for cautious and aggressive drivers. The vehicle at the first position of the platoon is at a constant velocity of 30 km/h in the beginning, then decelerates to 0 and stops for 10 second and then accelerates to 30 km/h again. For the cautious drivers' test scenario, the initial space is designed as 25 meters to ensure the vehicle is at a neutral state at the beginning. For the aggressive drivers' test, the initial space is 10 meters.



Fig 14. Shockwave with only cautious and aggressive drivers. [(a): Vehicle ID: 2582 cautious driver; (b) Vehicle ID: 2623 aggressive driver]

The propagations of the shockwave caused by the interruption of the first vehicle are plotted in Fig. 14. The results show that different congestion propagation patterns and shockwaves are reproduced for drivers with different neutral lines. It can be seen that the congestion is propagated at about 30 km/h in the traffic composed with aggressive drivers. In contrast, the speed of congestion propagation in traffic with cautious drivers is comparatively slower, which is only 17 km/h. As discussed in Section 4, the neutral line can determine the equilibrium relation

of the speed and space in NCF, so the impact of the neutral line on the congestion propagation can also be understood by NCF. The wave speed is determined by both the jam space and the slope of the space-speed relation, which is given by the following equation in NCF:

$$w = \omega/\tau \tag{11}$$

where  $\tau$  is the slope of the equilibrium spacing line and  $\omega$  is the jam space. In our test scenario, the jam space for the cautious driver is about 10 meters. For aggressive drivers, it is about 6 meters. The slope for the cautious driver is about 2 seconds, which is determined from the cautious driver's neutral line. For aggressive drivers, it is about 0.2 seconds as shown in Fig. 13. So applying Eq. (11), we can easily derive the wave speed for the cautious drivers to be approximately 18 km/h, and for the aggressive driver it is about 30 km/h. This matches the simulation results very well.

Some interesting phenomena is also observed in Fig. 14 such as the wave speed is not constant when the congestion propagating. In Laval's observations (Laval and Leclercq, 2010), the deviations from Newell's trajectory regarding the spacing can either be smaller or larger. When considering these factors, the congestion propagation will exhibit different propagation amplitude of time and space. Similarly, since the deviation is also implicitly integrated into (or learned by) our model from the real trajectory, it is not surprise that the congestion propagation speed is inconstant, which may imply some complex macroscopic traffic phenomena worthy of further investigation.

#### 5.3 Connection with the macroscopic hysteresis

The macroscopic hysteresis in the flow-density diagram is also one of the most prominent traffic phenomena. Analysis in this section attempts to connect the micro-hysteresis in asymmetric driving behavior to the macro-hysteresis in flow-density diagrams. We are interested in how the intensity of the asymmetric driving behavior and the corresponding micro-hysteresis affect the macro-level hysteresis in flow-density diagrams.

Equilibrium traffic has to be established as a benchmark for the macroscopic hysteresis. We applied the triangle fundamental diagram produced by NCF. The neutral line property extracted from individual trajectory is utilized to reproduce the equilibrium (Newell's) trajectory. Let the neutral line learned from the field trajectory data take the form of:

$$s = \tau v_f + \omega \tag{12}$$

Then the corresponding Newell's trajectory is:

$$\begin{cases} v_f(t) = v_l(t-\tau) \\ x_f(t) = x_l(t-\tau) - \omega \end{cases}$$
(13)

where  $v_f$  is the velocity of the following vehicle,  $v_l(t)$  is the velocity of the leading vehicle,  $x_l$  is the position of the leading vehicle, and  $x_f$  is the position of the following vehicle. The macroscopic traffic parameters are measured from the trajectories based on Laval's improved method, which uses the Edie's generalized definitions (Edie, 1961; Laval, 2011). In the equilibrium traffic, although no macro-hysteresis exists in density-flow diagrams, the micro-hysteresis still exists and can be observed. This is easy to understand since the driver has to take a relaxation time for returning to the equilibrium state when the leading vehicle decelerates or accelerates. That means a hysteresis at micro-level does not necessarily result in a macro-hysteresis in density-flow diagrams. We will refer this observed micro-hysteresis as the equilibrium asymmetry. However, when trajectories do not follow the Newell's trajectory, the micro-hysteresis will deviate from the equilibrium asymmetry. As pointed out by Laval and Leclercq (2010), two types of deviations from the Newell's trajectory can be identified in stop and go traffic from NGSIM data, which can be categorized as: a) the trajectory is above the Newell's trajectory, indicating a comparatively smaller spacing; b) the trajectory is below the Newell's trajectory, with a larger spacing. Examples of trajectories for these two groups are plotted in Fig. 15.

Obviously, deviations from the equilibrium trajectory will lead to different micro-hysteresis. We are interested in how the magnitude of these deviations in micro-hysteresis from the equilibrium asymmetry affects the macro-hysteresis in the density-flow diagram. For reproducing different magnitudes of the micro-hysteresis, stop-and-go traffic scenarios are simulated with different accelerating and decelerating intensity of the leading vehicle. The generated trajectories are measured to investigate the impact on the hysteresis in the density-flow diagram.



Fig 15. Trajectory samples comapred with Newell's trajectory

For the trajectory of which the deceleration path is closer to the leading vehicle compared with Newell's trajectory, the corresponding descending branch in the density-flow diagram is above the equilibrium path of the triangle fundamental diagram as shown in Fig.16, which indicates a higher density compared with the equilibrium. It can be seen that with an increase of the deceleration intensity of the leading vehicle, micro-hysteresis is more deviated from the equilibrium asymmetry, which results in a corresponding descending path of macro-hysteresis with a larger drift above the equilibrium path. This implies that larger deviation in micro-hysteresis will result in nonequilibrium states with larger departures from the equilibrium.



Fig 16. The micro- and macro-hysteresis for drivers following trajecotry category (a)-the deceleration branch is above Newell's trajectory



Fig 17. The micro- and macro-hysteresis for drivers following trajecotry category (b)-the deceleration branch is below Newell's trajectory

For the trajectory of which the deceleration path is farther from the leading vehicle compared with Newell's trajectory, the corresponding descending branch in MFD is below the equilibrium path as shown in Fig. 17, which indicates a lower density compared with the equilibrium. Similarly, as we can see from Fig. 17, if the macro-hysteresis is more deviated from the equilibrium asymmetry, the descending path in MFD shows a larger drift above the below path.

In summary, the observed equilibrium asymmetry in a space-speed diagram does not result in macro-hysteresis in the density-flow diagram. However, if the micro-hysteresis deviates from the equilibrium asymmetry, the macro-hysteresis in the density-flow diagram will be observed. The magnitudes of the deviations from the equilibrium states for micro- and macro-hysteresis are

positively related. Larger deviations from the equilibrium asymmetry in the micro-hysteresis will result in nonequilibrium traffic states with larger departures from the equilibrium.

#### 6. Conclusion

The asymmetric characteristics in driving behavior are analyzed using an SVR approach. Due to the outstanding learning ability of SVR, the asymmetric driving characteristics from the NGSIM data can be well tracked. New insights on asymmetric driving characteristics are gained through a comprehensive analysis from both micro- and macroscopic levels.

The intensity difference between acceleration and deceleration in driving behavior leads to a neutral line, which separates the speed-space diagram into acceleration and deceleration dominant areas. We found that the neutral line determines the tilt angle and position of the hysteresis in the speed-space diagram, while the leading vehicle's states affect the internal and external shapes of the hysteresis. The neutral line reflects the inherent driving behavior, while the states of the leading vehicles play the role of environmental stimuli.

At the macro-level, we found that the congestion propagates at different speeds in the traffic composed of drivers with different neutral lines. The relation between the neutral line and the wave-speed is found to be consistent with the NCF. Furthermore, new insights were also gained on how the micro- and macro-hysteresis are related. When the micro-hysteresis deviates from the equilibrium asymmetry, the macro-hysteresis in flow-density diagrams will be observed. The magnitudes of the deviations from the equilibrium states for micro- and macro-hysteresis are positively related. Larger deviations from the equilibrium asymmetry in the micro-hysteresis will result in nonequilibrium traffic states with larger departures from the equilibrium.

## Acknowledgements

We are thankful to the anonymous reviewers for contributing their time to review the manuscript and provide detailed comments which are very helpful in improving the quality of the paper.

#### References

- Ahn, S., Vadlamani, S., Laval, J., 2012. A method to account for non-steady state conditions in measuring traffic hysteresis. *Transportation Research Part C: Emerging Technologies*, In press.
- Aron, M., 1988. Car following in an urban network: simulation and experiments, 16th PTRC summer meeting, Sussex, UK.
- Bham, G.H., Benekohal, R.F., 2004. A high fidelity traffic simulation model based on cellular automata and car-following concepts. *Transportation Research Part C: Emerging Technologies* 12(1), 1-32.
- Brown, M.P., Grundy, W.N., Lin, D., Cristianini, N., Sugnet, C.W., Furey, T.S., Ares, M., Jr., Haussler, D., 2000. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proceedings of the National Academy of Sciences of the United States of America* 97(1), 262-267.
- Burbidge, R., Trotter, M., Buxton, B., Holden, S., 2001. Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Computers & chemistry* 26(1), 5-14.
- Chakroborty, P., Kikuchi, S., 1999. Evaluation of the General Motors based car-following models and a proposed fuzzy inference model. *Transportation Research Part C: Emerging Technologies* 7(4), 209-235.

- Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2(3), 1-27.
- Cherkassky, V., Ma, Y., 2004. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks* 17(1), 113-126.
- Edie, L., 1961. Car-Following and Steady-State Theory for Noncongested Traffic. *Operations Research* 9(1), 66-76.
- FHWA, 2007. NGSIM Next Generation Simulation, http://ngsim-community.org/.
- Gong, H., Liu, H., Wang, B.-H., 2008. An asymmetric full velocity difference car-following model. *Physica A: Statistical Mechanics and its Applications* 387(11), 2595-2602.
- Hair, J., Tatham, R., Anderson, R., Black, W., 1998. *Multivariate Data Analysis (5th Edition)*. Prentice Hall.
- Hastie, T., Tibshirani, R., Friedman, J., 2003. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
- Huang, C.-L., Chen, M.-C., Wang, C.-J., 2007. Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications* 33(4), 847-856.
- Johansson, G., Rummer, K., 1971. Drivers' brake reaction time. Human Factors 13(1), 1:23.
- Laval, J., Leclercq, L., 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. *Physical and Engineering Sciences* 368(1928), 4519-4541.
- Laval, J.A., 2011. Hysteresis in traffic flow revisited: An improved measurement method. *Transportation Research Part B: Methodological* 45(2), 385-391.
- Leutzbach, W., 1988. Introduction to the theory of traffic flow. Springer.
- Newell, G.F., 2002. A simplified car-following theory: a lower order model. *Transportation Research Part B: Methodological* 36(3), 195-205.
- Sakda, P., Hussein, D., 2007. Neural Agent Car-Following Models. Intelligent Transportation Systems, IEEE Transactions on 8(1), 60-70.
- Tordeux, A., Lassarre, S., Roussignol, M., 2010. An adaptive time gap car-following model. *Transportation Research Part B: Methodological* 44(8–9), 1115-1131.
- Vapnik, V., 1998. Statistical learning theory. Wiley.
- Vapnik, V.N., 1995. The nature of statistical learning theory. Springer-Verlag New York, Inc.
- Yeo, H., 2008. Asymmetric Microscopic Driving Behavior Theory. University of California Transportation Center.
- Yeo, H., Skabardonis, A., 2009. Understanding Stop-and-go Traffic in View of Asymmetric Traffic Theory. Transportation and Traffic Theory 2009: Golden Jubilee, in: Lam, W.H.K., Wong, S.C., Lo, H.K. (Eds.). Springer US, pp. 99-115.
- Zhang, H.M., 1999. A mathematical theory of traffic hysteresis. *Transportation Research Part B: Methodological* 33(1), 1-23.