2021 IEEE Intelligent Transportation Systems Conference (ITSC)
Indianapolis, USA. September 19-21, 2021

Connected Vehicle-Based Advanced Detection of "Slow-Down" Events on Freeways

Zhouqiao Zhao, *Student Member*, IEEE, Guoyuan Wu*, *Senior Member*, IEEE, Matthew J. Barth, *Fellow*, IEEE, Hossein Nourkhiz Mahjoub, Yasir Khudhair Al-Nadawi Laith Daman, and Shigenobu Saigusa

***Abstract*—From the perspective of an individual vehicle, the prediction of a "slow-down" or shockwave event on a freeway can help the driver reduce potential collision risks, enhance the driving experience, and reduce the cost of energy consumption and vehicle maintenance. From the perspective of traffic management, shockwave prediction may help regulate traffic flow effectively and allow for the response to (non-recurrent) incidents in a timely manner. In this paper, two real-time prediction algorithms are proposed and investigated, which are based on the high-resolution information provided from a set of connected vehicles within the communication range of the host vehicle. Both methods are able to predict the "slow-down" event under high traffic density at 3.51 seconds (on average) earlier than its occurrence. Both algorithm performances degrade with the decrease of the traffic density and penetration rate of the connected vehicles.**

I. Introduction

By accessing real-time information from multiple sources (e.g., other vehicles, road infrastructure, smart devices, and the Internet) via wireless communications, connected vehicles (CVs) can play a key role in addressing transportation-related socio-economic problems. From the perspective of individual drivers, CV technology is able to provide prompt downstream traffic information and customized driving guidance, which may go beyond the limitation (e.g., occlusion, spatial and temporal range) of on-board sensors such as radar, LiDAR, and cameras. Further, this downstream information may even encourage more cooperative maneuvers between vehicles [1]. Some studies focus on the provision of downstream aggregate traffic information (e.g., average lane speed) along freeways [2]. However, advanced detection of downstream traffic state transition such as formation and dissipation of traffic congestion (although more challenging) is of more value for the driver who can then anticipate a traffic state transition. This can help reduce potential collision risks, enhance driving experiences, and reduce costs in energy consumption and vehicle maintenance.

In this study, we develop and evaluate algorithms for the advanced detection of "slow-down" or shockwave events (from the perspective of the host vehicle), by leveraging CV technology. High-resolution information (such as location, speed) provided from a set of connected vehicles can be delivered to the host vehicle and used to develop driving guidance for the driver, to help him/her prepare for the downstream congestion formation. It is expected that the outcome of this exploratory study may lay the foundation for further developing and evaluating an advanced driving assistance system that may alert the driver about the upcoming deceleration or braking needs.

The major contributions of this research are that: 1) the proposed algorithms can predict the "slow-down" events based on current and historical data in real-time; 2) the prediction performance is investigated in different connected vehicle penetration rates and traffic conditions; and 3) we utilize microscopic traffic simulation to generate the data for a wide variety of traffic conditions, providing a robust solution.

The remainder of this report is organized as follows: *Section II* reviews background information on topics related to this project, including probe vehicle or connected vehicle-based traffic state estimation as well as traffic shockwave analysis and detection. *Section III* examines the problem formulation for advanced detection of a "slow-down" event, followed by the presentation of two proposed algorithms, i.e., cell-based average deceleration (CAD) and speed standard deviation (SSD) in *Section IV*, to detect "slow-down" events. *Section V* describes the evaluation methodology and performance metrics, and the setup of the simulation network in PTV VISSIM as well as simulation results are presented in *Section VI*. The last section summarizes the findings from the project and discusses a few key directions for further exploration in the future.

II. Background

A. Vehicle as a Sensor (VaaS)

The majority of existing studies on detecting traffic congestion rely on traditional infrastructure-based sensors such as inductive loop detectors [3] and cameras [4], which generally provide aggregate traffic flow information (i.e., traffic volume, average speed) or even trajectories of all vehicles within limited road segments. However, information from these sensors is restricted in space, and therefore does not provide a full sketch of the spatial-temporal traffic dynamics along specific routes. With the advent of probe vehicles and connected vehicles (CVs), or Vehicle as a Sensor (VaaS) in a broader sense, traffic or vehicle state detection based on (communication-enabled) probe vehicles (or "floating cars") continues to be a topic

Z. Zhao, G, Wu, and M. J Barth are with the Bourns College of Engineering, Center for Environmental Research and Technology (CE-CERT), University of California, Riverside, CA, 92507, USA (e-mail: zzhao084@ucr.edu; gywu@cert.ucr.edu; barth@ece.ucr.edu).
H. Mahjoub, Y. Al-Nadawi, L. Daman, S. Saigusa are with Honda Motor Co., Ltd. (e-mail: hnourkhizmahjoub@Honda-ri.com yalnadawi@Honda-ri.com; ldaman@Honda-ri.com shigenobu_saigusa@jp.honda)

2021 IEEE International Intelligent Transportation Systems Conference (ITSC) | 978-1-7281-9142-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/ITSC48978.2021.9564594

of great interest in the research community. However, most of these VaaS related studies have focused on: 1) the macroscopic or mesoscopic traffic state estimation (such as travel time, queue length), given the limited penetration rate of CVs or by integrating the measurement from other sources (i.e., infrastructure-based sensors) [5]–[8]; and 2) individual vehicle's abnormal state identification and broadcast (via communication) to enable customized driving guidance for the drivers in host vehicles [9]–[11]. It is much more challenging to extract the information from these "mobile" sensors to reconstruct detailed and transient dynamics of traffic flow, especially if the penetration rate of CVs is low.

B. *Traffic Shockwave Detection and Analysis*

From the traffic perspective, a shockwave is highly related to the propagation of traffic congestion upstream, thus potentially resulting in deceleration and braking maneuvers (or "slow-down" events) of the host vehicle that approaches the congestion area. By definition, a shockwave is a boundary in a traffic stream that represents a discontinuity in the flow-density domain [12]. In particular, the shockwave formation can be considered as a consistent pattern of the significant reduction in speed propagating over space and time. Shockwave analysis is a technique to identify traffic congestion along a roadway and to estimate the rate of shockwave formation and dissipation. Some researchers investigated real-world naturalistic driving data from HD video cameras mounted on the pole (e.g., Next Generation SIMulation [13]) or drones [14] to analyze shockwave characteristics and dynamics. For example, Lu and Skabardonis [15] applied a smoothing filter to the vehicle trajectory data collected on the US-101 site (one of the four NGSIM datasets) for damping noises in raw data, and performed analyses (e.g., speed profiles and time-space diagrams) to visualize the shockwave formation and dissipation on a lane basis. Based on the analysis, they proposed an algorithm to estimate the propagation speed of shockwave on freeways. Other than trading the trajectory data as the signal, recently, researchers such as use machine learning (ML) approaches to predict traffic flow breakdown.

Although the real-world datasets can provide realistic detailed driving data for the entire traffic flow along roadway segments (with some measurement errors [16]), the information is only applicable to specific scenarios (i.e., within limited spatio-temporal regions and under certain traffic conditions). To facilitate the exploratory study, such as the determination of model parameters and algorithm thresholds, microscopic traffic simulation is a better fit for the scope, provided the simulation model has been calibrated with real-world data.

III. Problem Formulation

The objective of this study is to explore effective algorithms for the advanced detection of slow-down events, by taking full advantage of vehicle-to-vehicle (V2V) communications. To formulate the problem, a host vehicle's bird-eye view associated with traveling along the same lane over consecutive time windows can be stitched together, as illustrated in Fig. 1 (the red trajectory is the host vehicle). In this space-time diagram, the red solid curve represents the (spatio-temporal) trajectory of the host vehicle, and the yellow dashed curve represents the maximum (downstream) communication range of the host vehicle along its trajectory. The shaded area under the yellow dashed curve, in principle, denotes the maximum spatio-temporal region (or "information availability region") where all the possible downstream traffic information can be available to the host vehicle up to the time instant to make real-time decision. Based on this information, the host vehicle might be able to detect imminent downstream shockwaves that may result in its "slow-down" maneuvers. Interestingly, if the host vehicle travels fast (e.g., up to time t'), the information availability region (IAR) is small or the maximum backtracking duration is short. If the host vehicle speed is slow (e.g., between time t' and t), the IAR is large or the maximum backtracking duration is long. In addition, if the roadway is (spatially) divided into cells and the backtracking time window is selected, then the area of information availability region that can be covered by the horizontal bars will depend on the size of the cell and the length of backtracking time window. By varying these parameters (i.e., cell size and window length) with the host vehicle speed, better coverage of the IAR would be expected and potentially more information would be available for the host vehicle to support decision making in real-time.

In this paper, we first determine if the "slow-down" event occurs (from the host vehicle's point of view), then develop the algorithms to extract features and characterize traffic states, and finally examine if the characterization results may provide indication of the "slow-down" occurrence. Once we determine if a "slow-down" event happens, we will check backward in time if there are any features or signs from the traffic along the respective road segment (on a cell basis) that can be utilized to predict a "slow-down" for the host vehicle.

Some existing studies have focused on detecting abnormal driving activities or events (such as harsh braking,

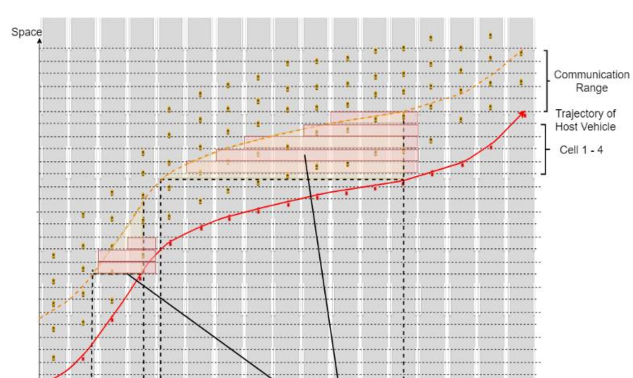

Fig. 1. Stitched bird-eye views with respect to the host vehicle in a space-time diagram.

aggressive swerving) based on analyzing speed trajectories, acceleration, and heading information that may be available from smartphones [17] or on-board GPS/IMU [18]. These studies have applied a set of rules, fuzzy logic, or machine learning algorithms to identify these events. As an exploratory approach, we propose a rule-based algorithm to determine the "slow-down" event in this study.

More specifically, we use the acceleration profile to characterize a "slow-down" event and determine its onset and duration, as shown in Fig. 2. The proposed "slow-down" identification method has two steps: 1) to find successive deceleration sequence that satisfies both maximum deceleration and duration conditions, including: a) deceleration is less than -2 m/s^2 and lasts for at least 2 seconds; b) deceleration is between -2 m/s^2 and -1 m/s^2 and lasts for at least 4 seconds; and c) deceleration is between -1 m/s^2 and -0.5 m/s^2 and lasts for at least 6 seconds; and 2) to group successive deceleration sequences with small enough time interval into one event.

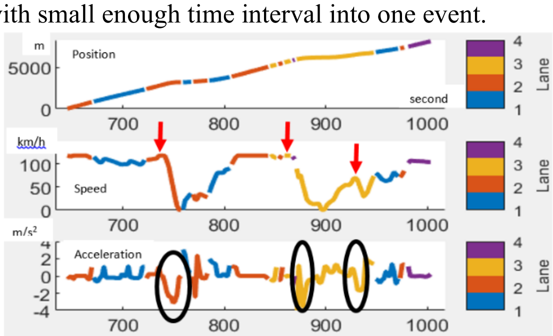

Fig. 2. An example showing the characterization of slow-down events.

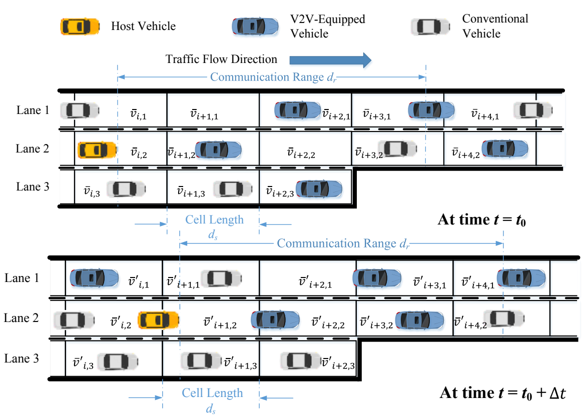

Fig. 3. Illustration of cell-based average deceleration.

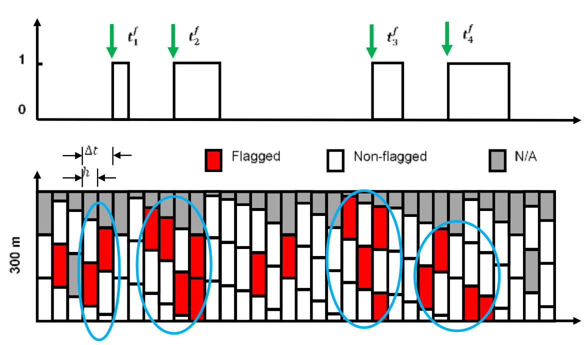

Fig. 4. Illustration of flag raise-up.

IV. Description of Algorithms

A. Algorithm 1: Cell-based Average Deceleration

The first proposed metric to identify slow-down events via connected vehicles is called Cell-based Average Deceleration (CAD). As shown in Fig. 3, the entire road network is first discretized into "cells" (whose size, d_s, could be fixed or scenario dependent due to different traffic demands) on a lane-by-lane basis. A host vehicle continuously monitors its downstream traffic state at the cell level within a specific communication range, d_r (e.g., 300 meters), calculating the average deceleration for each individual cell of interest according to the following equation:

$$\delta_{cell(i,j)}(t;\Delta t)=\frac{\bar{v}_{cell(i,j)}(t)-\bar{v}_{cell(i,j)}(t-\Delta t)}{\Delta t} \quad (1)$$

and,

$$\bar{v}_{cell(i,j)}(t) = \begin{cases} \frac{\sum_{k\in K_{cell(i,j)}(t)} v_k(t)}{N_{cell(i,j)}(t)}, & K_{cell(i,j)}(t) \neq \emptyset \\ v^{limit}, & otherwise \end{cases} \quad (2)$$

where $\delta_{cell(i,j)}(t;\Delta t)$ denotes the average deceleration (or speed gradient over time) for $cell(i,j)$ at time t; i represents the index of segment (i.e., a group of cells across lanes); j represent the lane index; $\bar{v}_{cell(i,j)}(t)$ is the average instantaneous speed of vehicles (of interest) whose front bumper are within the boundary of $cell(i,j)$ or simply termed as "in $cell(i,j)$" at time t; Δt is a user-defined look-back time step (which could be fixed or scenario dependent due to different traffic demands); $K_{cell(i,j)}(t)$ denotes the index set of vehicles (of interest) in $cell(i,j)$ at time t; k is the vehicle index in the set $K_{cell(i,j)}(t)$; $v_k(t)$ is the instantaneous speed of vehicle k at time t; and $N_{cell(i,j)}(t)$ represents the number of vehicles (or the size of $K_{cell(i,j)}(t)$) in $cell(i,j)$ at time t. Please note that if there is no vehicle in $cell(i,j)$ at time t (i.e., $K_{cell(i,j)}(t) = \emptyset$), then $\bar{v}_{cell(i,j)}(t)$ is assumed to be the roadway speed limit v^{limit} or other reasonable estimation (e.g., $\bar{v}_{cell(i,j)}(t-h)$ where h is the measurement update interval). Also, only the cells within the range of communication with respect to the host vehicle, d_r (e.g., 300 meters) will be considered in the calculation.

In addition, to avoid any noise or fluctuations in measuring instantaneous speed, a user-defined time window $\tau \triangleq n \cdot h$ is applied to consecutive measurement update intervals (e.g., 0.1 seconds for the DSRC-enabled connected vehicle environment) as a filter mechanism, where n is a positive integer. Then, the slow-down events, on a cell basis would be identified and the associated flags would be raised for each respective cell according to the conditions below:

$$\delta_{cell(i,j)}(t-c_2\cdot h;\Delta t) \leq \theta < 0, \forall c_2 = 0,1,\cdots,n.$$

where θ is user-defined thresholds of cell-based average deceleration to determine the slow-down events; and n is user-defined positive integer to determine the sizes of filtering windows for the slow-down events, respectively.

It should be noted that in this phase, we mainly focus on Type I. There are three major steps in this process:

a) Raise flag if any of the valid cell satisfies the conditions (as described above) for successive h steps. More specifically, at each time step, the host vehicle may have the prediction results from multiple cells where the flag is raised if the predefined threshold(s) are satisfied. Then, the results of all cells are combined using OR logic. The illustration of a "flag raise-up" is shown in Fig. 4. Since the prediction signal calculated from CAD (similar to Approach II – SSD) can be fluctuant due to the randomness of the driving behaviors, change of valid cells, and change of vehicles in the cells, it is necessary to smooth the signals and make them robust enough for performance evaluation.

b) Group any successive prediction sequences with too small-time interval(s) into one detection; and

c) Remove any flag raising event that has too short duration.

The detailed algorithm is shown in the following pseudo-code:

Algorithm 1: Cell-based Average Deceleration

```
Initialize CAD_combine, CAD_prediction, CAD_start, CAD_end
FOR each step in the vehicle trajectory
  CAD_prediction_lane = 0
  FOR each valid cell from different lanes at current step
    IF CAD_veh[current lane][current step] <= Threshold
      CAD_prediction_lane = 1
    ENDIF
    CAD_combine += CAD_prediction_lane
  ENDFOR
  IF CAD_combine > 0 and CAD_prediction[last h steps] == True
    CAD_prediction[current step] = True
  ENDIF
ENDFOR
FOR each step in the vehicle trajectory
  IF CAD_prediction[last step] == 1 and CAD_prediction[current step] == 0
    CAD_end = current step
  ELSEIF CAD_prediction[last step] == 0 and CAD_prediction[current step] == 1 and current step – Slow_down_end <= Threshold
    CAD_prediction[Slow_down_end : current step] = True
  ENDIF
ENDFOR
FOR each step in the vehicle trajectory
  IF CAD_prediction[last step] == 0 and CAD_prediction[current step] == 1
    CAD_start = current step
  ELSEIF CAD_prediction[last step] == 1 and CAD_prediction[current step] == 0   and current step – CAD_start <= Threshold
    CAD_prediction[CAD_start : current step] = False
  ENDIF
ENDFOR
```

B. *Algorithm 2: Speed Standard Deviation (SSD)*

This proposed approach is enlightened by Elfar et al. [19] which utilizes the vehicle distribution information to identify the slow-down events along a road segment. We have modified this method for real-time slow-down event prediction purpose with the following four steps:

a) Divide the road segment into cells with a user-defined length (e.g., 200 ft as suggested in the study). This step is similar to some procedure in Approach 1.

b) Calculate the average speed of each vehicle along each road segment at the lane level over a user-defined time interval, Δt:

$$\bar{v}_{k,cell(i,j)}(t,\Delta t) = \frac{\sum_{A_k} v_k(t,\Delta t)}{|A_k|} \tag{3}$$

where $v_k(t)$ represents the instantaneous speed for vehicle k; A_k denotes the set of samples where vehicle k travels in $cell(i,j)$ from the time $t-\Delta t$ to t; and $|\cdot|$ is the cardinality of a set. To have an enough number of the valid cells for prediction, the time interval, Δt, should be much shorter than 10 seconds. As a result, the cell size should be adjusted accordingly to have enough vehicles in each cell.

c) Calculate the speed standard deviation (SSD) of the average speeds of all individual vehicles for each cell at each time step that are obtained from Step b) using

$$SSD_{cell(i,j)}(t,\Delta t) = \sqrt{\frac{\sum_{k\in K_{cell(i,j)}}\left(\bar{v}_{k,cell(i,j)}(t,\Delta t)-\bar{v}_{cell(i,j)}(t,\Delta t)\right)^2}{N_{cell(i,j)}}} \tag{4}$$

$$\bar{v}_{cell(i,j)}(t,\Delta t) = \frac{\sum_{k\in K_{cell(i,j)}} \bar{v}_{k,cell(i,j)}(t,\Delta t)}{N_{cell(i,j)}} \tag{5}$$

d) Raise flag when SSD satisfy the following condition:

$$SSD_{cell(i,j)}(t-c_3\cdot h;\Delta t) > \theta_2, \forall c_3 = 0,1,\cdots,n_3$$

Similar to *Algorithm 1*, there are also three major steps, as shown below, representing raising up of flags, fusion of neighbor events, and removal of short events.

Algorithm 2: Speed Standard Deviation

```
Initialize SSD_combine, SSD_prediction, SSD_start, SSD_end
FOR each step in the vehicle trajectory
  SSD_prediction_lane = 0
  FOR each valid cell from different lanes at current step
    IF SSD_veh[current lane][current step] >= Threshold
      SSD_prediction_lane = 1
    ENDIF
    SSD_combine += SSD_prediction_lane
  ENDFOR
  IF SSD_combine > 0 and SSD_prediction[last h steps] == True
    SSD_prediction[current step] = True
  ENDIF
ENDFOR
FOR each step in the vehicle trajectory
  IF SSD_prediction[last step] == 1 and SSD_prediction[current step] == 0
    SSD_end = current step
  ELSEIF SSD_prediction[last step] == 0 and SSD_prediction[current step] == 1 and current step – Slow_down_end <= Threshold
    SSD_prediction[Slow_down_end : current step] = True
  ENDIF
ENDFOR
FOR each step in the vehicle trajectory
  IF SSD_prediction[last step] == 0 and SSD_prediction[current step] == 1
    SSD_start = current step
  ELSEIF SSD_prediction[last step] == 1 and SSD_prediction[current step] == 0   and current step – SSD_start <= Threshold
    SSD_prediction[SSD_start : current step] = False
  ENDIF
ENDFOR
```

V. Evaluation Methodology

In this study, we use Confusion Matrix (CM) for performance evaluation. CM also known as error matrix is a specific table layout that allows presentation of the performance measurement of a classification algorithm [20]. As shown in Fig. 5, each row of the matrix represents the instances in a predicted class, while each column represents the instances in an actual class. The four entries of the CM are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

For event-based performance analysis, the matrix can be fill out by simply counting the number of each event. In this study, however, to measure the performance of the proposed "slow-down" prediction methods, the definition of predicted class needs to be modified, because the "non-slow-down" incidents are not counted. Therefore, we use "rate" rather than "number" to quantify the predicted class. Below is a more detailed description.

- True Positive (TP) event is defined as the earliest prediction signal (i.e., flag raise-up signal) within a time window (TW) before the onset of the associated actual "slow-down" event;
- True Positive Rate (TPR) equals to the number of TP events divided by the total number of "slow-down" events;
- False Negative (FN) event is defined if no prediction is made within a time window (TW) before the onset of the associated actual "slow-down" event;
- False Negative Rate (FNR) equals to the number of FN divided by the total number of "slow-down" events;
- False Positive Rate (FPR) is defined as the time duration of flag raise-up outside the time window (TW) divided by overall "non-slow-down" time duration;

True Negative Rate (TNR) is defined as the time duration without raising flag outside the time window (TW) divided by overall "non-slow-down" time duration. Fig. 6. illustrates definitions of the above terms. In addition, the earliest time (within the predefined time window) to detect a "slow-down" in advance is also an important performance metric. The time from the earliest flag raise-up of a TP event to the actual start time of a "slow-down" event is

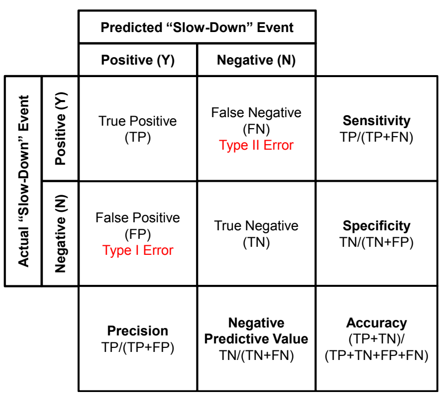

Fig. 5. Illustration of Confusion Matrix.

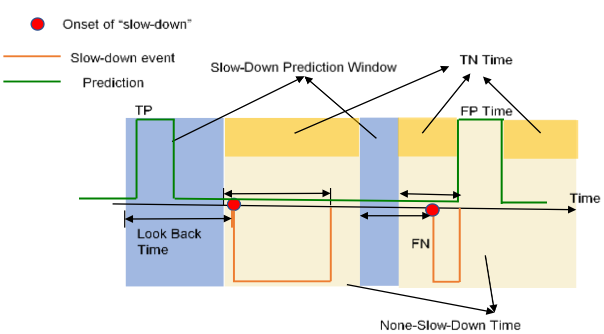

Fig. 6. Definitions of modified predicted class in CM.

called detection time. The average detection time denotes the average of the detection time for each TP of the host vehicle.

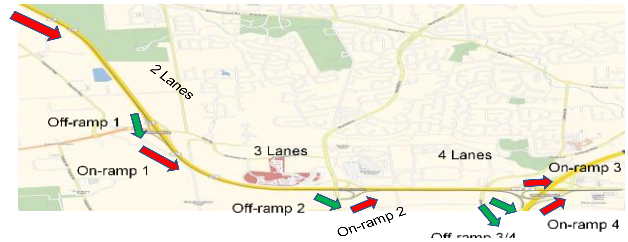

Fig. 7. PTV VISSIM network: US-33 E.

VI. Preliminary Simulation Study

A. Simulation Setup

Over the performance period of this project, hawse have coded a real-world network in PTV VISSIM, a microscopic traffic simulation platform, with up-to-date roadway geometry as shown in Fig. 7. More specifically, the network is a selected segment of US-33 E in Ohio which is a five-mile-long road stretch with two to four lanes. The off-ramp 3 and off-ramp 4 are the route to I-270 N and I-270 S, respectively. Between on-ramp 2 and off-ramp 3 and off-ramp 4, there are 4 lanes in total, and the traffic volume is very high. Also, when the traffic volume from upstream is high, the on-ramp 1 would be another bottleneck area where the merging vehicles from on-ramp 1 slow down the mainline traffic in both left and right lanes.

B. Traffic Demand

The input volume (which is considered as a moderate demand scenario) and the static vehicle routing decision have been carefully calibrated in PTV VISSIM micro-simulation implementation. To investigate how the prediction performance of the proposed methods would vary in different traffic demands (or potentially different levels of service along road segments), we select another traffic input volume (distribution) as a high demand scenario. The volume of each tested scenario is shown in TABLE I.

TABLE I. Set-up of Traffic Demands in Simulation

Volume (veh/h) / Scenario	US33 Upstream	On-ramp 1	On-ramp 2	On-ramp 3	On-ramp 4
Moderate Demand (Calibrated)	1794	800	1370	348	327
High Demand	2500	1500	2000	348	327

C. *Artificial Slow-Downs*

To have more speed variation samples, during the 30-minute simulation, four artificial slow-down incidents have been generated using the COM interface. Other than control speeds of the slow-down vehicles directly, the research team codes a script using the COM interface to change the desired speed of the target vehicles to achieve more realistic incidents. At the simulation time 600s, 800s, 1000s, and 1200s, the target vehicles in the segment between on-ramp 2 and on-ramp 3 are selected randomly, and their desired speeds are set to be 5 km/h for 60 seconds. As a result, the target vehicles slow down smoothly governed by the driving model in PTV VISSIM, and other vehicles respond the target vehicles accordingly.

D. *Algorithm Parameter Selection*

The research team conducts a 30-minute simulation for each scenario. The vehicles' trajectories, the cell-based spatio-temporal heat-map can be generated offline. The trajectory of the host vehicle is color-coded based on the lane information. All the vehicles operating in the network over five minutes are used as potential host vehicles for prediction analysis. Then, the average performance measurements can be calculated by weighting with the respective operating time.

Besides the conditions to identify the slow-down event, the selection of algorithm parameters may significantly impact the system performance. The choice of these parameters may balance the sensitivity, accuracy, and robustness of the algorithm. In this study, we apply a trial-and-error approach based on observations from the simulation results. The parameters used for this study are shown in TABLE II.

TABLE II. PARAMETERS SELECTION

	Δt	Threshold	Successive Step	Grouping interval	Deactivation duration
CAD	1 s	-5 m/s^2	5 steps	5 s	5 s
SSD	1 s	10 m/s	5 steps	5 s	5 s

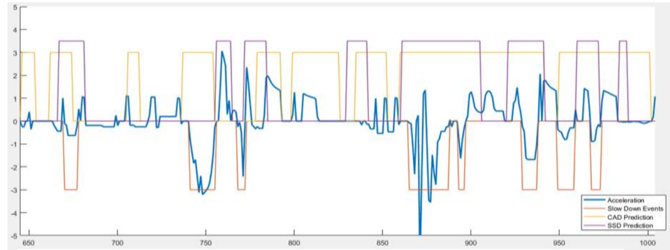

Fig. 8. An example of prediction result

E. *Result Analysis*

In this section, we present some preliminary results from the simulation study. Fig. 8. presents a typical example of prediction result. The blue curve denotes the acceleration of the host vehicle; the orange curve denotes the slow-down events; the yellow curve denotes the CAD prediction; and the purple curve denotes the SSD prediction. As can be observed from the figure, CAD approach may predict most of the slow-down events earlier than SSD approach, but it also has a relatively higher false positive rate. TABLE III summarizes the overall simulation results, which are in line with the observation from the above example. Regarding different traffic demands and penetration rates of connected vehicles, the proposed algorithms' performance may be degraded in moderate demands and low penetration rates, mainly due to the lack of available information for robust prediction of "slow-down" events.

TABLE III. SET-UP OF TRAFFIC DEMANDS IN SIMULATION

Scenario	Approach	Avg. Travel Time (s)	Avg. # of Events	Avg. Detection Time (s)	TPR	FNR	FPR	TNR
High Demand 100% CV	CAD	329.58	5.77	3.51	0.70 (1.4%)	0.30 (-3.2 %)	0.35 (1.7%)	0.65 (-7.1%)
	SSD			3.26	0.69	0.31	0.30	0.70
High Demand 75% CV	CAD	330.18	5.86	3.28	0.65 (6.5%)	0.35 (-10.2%)	0.30 (30.4%)	0.70 (-9.1%)
	SSD			2.87	0.61	0.39	0.23	0.77
High Demand 50% CV	CAD	328.42	5.71	3.55	0.67 (15.5%)	0.33 (-21.4%)	0.23 (27.8%)	0.77 (-6.1%)
	SSD			2.78	0.58	0.42	0.18	0.82
High Demand 25% CV	CAD	330.17	6.12	2.94	0.50 (38.9%)	0.50 (-21.9%)	0.11 (57.1%)	0.89 (-4.3%)
	SSD			1.74	0.36	0.64	0.07	0.93
Moderate Demand 100% CV	CAD	261.89	2.51	2.47	0.44 (63.0%)	0.56 (-23.3%)	0.15 (114.3%)	0.85 (-8.6%)
	SSD			1.84	0.27	0.73	0.07	0.93
Moderate Demand 75% CV	CAD	260.83	2.33	2.62	0.57 (54.0%)	0.43 (-31.7%)	0.12 (100.0%)	0.88 (-6.4%)
	SSD			1.66	0.370	0.63	0.06	0.94
Moderate Demand 50% CV	CAD	259.45	2.24	1.71	0.28 (133.3%)	0.72 (-18.1%)	0.05 (150.0%)	0.95 (-3.1%)
	SSD			0.37	0.12	0.88	0.02	0.98
Moderate Demand 25% CV	CAD	258.80	2.73	1.00	0.17 (41.7%)	0.83 (-5.7%)	0.01 (0%)	0.99 (0%)
	SSD			0.94	0.12	0.88	0.01	0.99

VII. Conclusions and Future Directions

A. Key Findings

At 100% CV penetration rate, cell-based average deceleration (CAD) approach and speed standard deviation (SSD) have very comparative performance, where the accuracy is about 70%. With a decrease of the CV penetration rate, the prediction performance of both CAD and SSD decreases accordingly due to less information being available. Both CAD and SSD become less sensitive to the "slow-down" events, which can be verified by the decreasing trend of True Positive Rate (TPR) and the increasing trend of False Positive Rate (FPR). When the CV penetration rate is as low as 25%, the TPRs for both CAD and SSD are less than or equal to 0.5, which makes the prediction unreliable. Compared with SSD, CAD is more sensitive because it consistently has larger TPRs over different scenarios. However, due to its higher sensitivity, CAD also has more FPRs, which reduces the prediction accuracy. In terms of the detection time, CAD typically has an earlier prediction compared to SSD. When the traffic becomes less congested (e.g., only 1.67 slow-down events per trip on average), the number of vehicles in the valid cell and even the number of valid cells decrease significantly. Therefore, the prediction performance at moderate demand is worse than high demand with the same CV penetration rate. In particular, CAD is much more reliable than SSD during the moderate traffic condition. It should be noted that all the conclusions above are based on the simulated scenarios and the robustness of these findings needs further investigation.

B. Future Directions

The proposed approaches are governed by a number of predefined thresholds and parameters, which can largely impact the performance of the prediction. In this current project, such parameters are chosen based on the observations from a limited set of collected data. Therefore, one potential research direction is to find a systematic way and leverage the power of data-driven methods (e.g., machine learning technique), to find the optimal set of parameters or the best schema to select these parameters. It is admitted that a single set of parameters may not be able to cover all the traffic scenarios. For example, the optimal CAD thresholds and cell size in low-speed scenarios might be different from those in high-speed scenarios, as explained in the section of A "Big Picture". Therefore, adaptive method(s) that can adjust the parameters according to the contemporary situations is worth further investigation. Finally, end-to-end learning techniques using deep neural networks (DNNs) is a promising candidate method for further exploration. This may help not only the identification of "ground truth" (i.e., slow-down) events, but also the detection/prediction of these events. By leveraging the microscopic traffic simulation software such as PTV VISSIM, a large amount of data can be easily generated for the learning purpose. Nevertheless, how to transfer the learning results from simulation environment to the real-world situation is an interesting topic.

Acknowledgment

This research was funded by the Honda Motor Co., Ltd. under the Agreement 4500328057. The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views of the Honda Motor Co., Ltd.

References

[1] D. Tian, G. Wu, P. Hao, K. Boriboonsomsin, and M. J. Barth, "Connected vehicle-based lane selection assistance application," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 7, pp. 2630–2643, 2018.

[2] D. Tian *et al.*, "Evaluating the effectiveness of V2V-based lane speed monitoring application: a simulation study," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 1592–1597.

[3] J. Guerrero-Ibáñez, S. Zeadally, and J. Contreras-Castillo, "Sensor technologies for intelligent transportation systems," *Sensors*, vol. 18, no. 4, p. 1212, 2018.

[4] P. Chakraborty, Y. O. Adu-Gyamfi, S. Poddar, V. Ahsani, A. Sharma, and S. Sarkar, "Traffic congestion detection from camera images using deep convolution neural networks," *Transp. Res. Rec.*, vol. 2672, no. 45, pp. 222–231, 2018.

[5] J. C. Herrera and A. M. Bayen, "Incorporation of Lagrangian measurements in freeway traffic state estimation," *Transp. Res. Part B Methodol.*, vol. 44, no. 4, pp. 460–481, 2010.

[6] N. Bekiaris-Liberis, C. Roncoli, and M. Papageorgiou, "Highway traffic state estimation with mixed connected and conventional vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3484–3497, 2016, doi: 10.1109/TITS.2016.2552639.

[7] S. M. Khan, K. C. Dey, and M. Chowdhury, "Real-time traffic state estimation with connected vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1687–1699, 2017.

[8] L. Li, R. Jiang, Z. He, X. M. Chen, and X. Zhou, "Trajectory data-based traffic flow studies: A revisit," *Transp. Res. Part C Emerg. Technol.*, vol. 114, pp. 225–240, 2020.

[9] J. Ferreira *et al.*, "Driver behavior profiling: An investigation with different smartphone sensors and machine learning," *PLoS One*, vol. 12, no. 4, p. e0174959, 2017.

[10] J. Yu, Z. Chen, Y. Zhu, Y. Chen, L. Kong, and M. Li, "Fine-grained abnormal driving behaviors detection and identification with smartphones," *IEEE Trans. Mob. Comput.*, vol. 16, no. 8, pp. 2198–2212, 2016.

[11] M. A. Abdel-Aty *et al.*, "Using Smartphone as On-board unit (OBU) Emulator Implementation Study," 2020.

[12] A. D. May, *Traffic flow fundamentals*. 1990.

[13] Federal Highway Administration, "Next Generation SIMulation (NGSIM)," 2021. https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm (accessed Jan. 07, 2021).

[14] E. Barmpounakis and N. Geroliminis, "On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment," *Transp. Res. part C Emerg. Technol.*, vol. 111, pp. 50–71, 2020.

[15] X.-Y. Lu and A. Skabardonis, "Freeway traffic shockwave analysis: exploring the NGSIM trajectory data," 2007.

[16] B. Coifman and L. Li, "A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset," *Transp. Res. Part B Methodol.*, vol. 105, pp. 362–377, 2017.

[17] D.-N. Lu, D.-N. Nguyen, T.-H. Nguyen, and H.-N. Nguyen, "Vehicle mode and driving activity detection based on analyzing sensor data of smartphones," *Sensors*, vol. 18, no. 4, p. 1036, 2018.

[18] H. Li, H. Wang, L. Liu, and M. Gruteser, "Automatic unusual driving event identification for dependable self-driving," in *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, 2018, pp. 15–27.

[19] A. Elfar, C. Xavier, A. Talebpour, and H. S. Mahmassani, "Traffic shockwave detection in a connected environment using the speed distribution of individual vehicles," *Transp. Res. Rec.*, vol. 2672, no. 20, pp. 203–214, 2018.

[20] S. V Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sens. Environ.*, vol. 62, no. 1, pp. 77–89, 1997.