LSTM-based Human-Driven Vehicle Longitudinal Trajectory Prediction in a Connected and Autonomous Vehicle Environment

Lei Lin, Ph.D.
Senior Research Scientist
Goergen Institute for Data Science
University of Rochester, Rochester, NY, 14620
Email: Lei.Lin@rochester.edu

Siyuan Gong, Ph.D.
Associate Professor
School of Information Engineering
Chang’an University, Xi’an, Shanxi, China
Email: sgong@chd.edu.cn

Srinivas Peeta, Ph.D. *
Frederick R. Dickerson Chair & Professor
School of Civil and Environmental Engineering
H. Milton Stewart School of Industrial Engineering
Georgia Institute of Technology, Atlanta, GA, 30332
Tel: 404-894-2243
Email: peeta@gatech.edu

Xia Wu
Research Assistant
School of Information Engineering
Chang’an University, Xi’an, Shanxi, China
Tel: +86 02962630049
Email: viichelle_wu@chd.edu.cn

* Corresponding Author
ABSTRACT

The advent of connected and autonomous vehicles (CAVs) will change driving behavior and travel environment, and provide opportunities for safer, smoother, and smarter road transportation. During the transition from the current human-driven vehicles (HDVs) to a fully CAV traffic environment, the road traffic will consist of a “mixed” traffic flow of HDVs and CAVs. Equipped with multiple sensors and vehicle-to-vehicle communications, a CAV can track surrounding HDVs and receive trajectory data of other CAVs in communication range. These trajectory data can be leveraged with recent advances in deep learning methods to potentially predict the trajectories of a target HDV. Based on these predictions, CAVs can react to circumvent or mitigate traffic flow oscillations and accidents. This study develops attention-based Long Short-Term Memory (LSTM) models for HDV longitudinal trajectory prediction in a mixed flow environment. The model and a few other LSTM variants are tested on the Next Generation SIMulation (NGSIM) US 101 dataset with different CAV market penetration rates (MPRs). Results illustrate that LSTM models that utilize historical trajectories from surrounding CAVs perform much better than those that ignore information even when the MPR is as low as 0.2. The attention-based LSTM models can provide more accurate multi-step longitudinal trajectory predictions. Further, we conduct grid-level average attention weight analysis and identify the CAVs with higher impact on the target HDV’s future trajectories.

Keywords: Longitudinal Trajectory Prediction, Long Short-term Memory Model, Attention Weight, Connected and Autonomous Vehicle, Human-driven Vehicle, Market Penetration Rate
INTRODUCTION

In the past few decades, detailed vehicle trajectory data have been collected through advances in technologies such as Global Position System (GPS) (1, 2) and high-resolution digital images from camera (3), helicopter (4) or Unmanned Aerial Vehicle (5). With the accumulation of these data, vehicle trajectory prediction has attracted considerable interest from the perspectives of safety and efficiency of the road transportation system.

Vehicle trajectory prediction can be beneficial in many scenarios. As an example, for a vehicle equipped with Advanced Driver Assistance Systems (ADAS) such as Adaptive Cruise Control, Collision Warning System, and Emergency Braking System, predicting future trajectories of vehicles detected in its vicinity can alert the driver about certain hazard scenarios, and hence can reduce the potential traffic accident risk (6–8). As another example, modeling the vehicle’s own trajectories can help the ADAS to create specific driving profiles and this trajectory prediction can improve the driving comfort for turn-making tasks (9).

More recently, self-driving or autonomous vehicle (AV) technologies are being explored to address multiple travel-related objectives. An AV is equipped with various powerful sensors like camera, Lidar, Radar, GPS, and ultrasonic to detect and perceive its surrounding environment. Based on the levels of autonomy defined by the SAE International’s standard J3016 (10), a fully-autonomous vehicle can operate on any road under any condition. An AV has the potential to change driving behavior and travel environment, providing opportunities for safer, smoother, and smarter road transportation. However, the fatal accident between a pedestrian and Uber’s self-driving vehicle in March, 2018 illustrates that there is significant room to improve AV technologies, and safety should always be considered with the highest priority in this process (11).

Connected vehicle (CV) technologies are also being deployed to improve the safety and mobility of our transportation system by enhancing situational awareness and traffic state estimation through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. A CV can communicate with other CVs within approximately 300 meters to sense what other travelers are doing (12). V2V and V2I can enable applications like cooperative collision warning, providing traffic signal status information in real time, and so on (13). CV applications require low latency and high-reliability communications; hence, efficient CV data capture and transmittal are of paramount importance (14, 15).

Although the CV and AV technological advances have been promoted by different stakeholders and follow independent paths, a strong synergy exists between them (13). Innovative opportunities arise through the simultaneous utilization of connectivity and automation. A connected and autonomous vehicle (CAV) can leverage automation functions while communicating with other CAVs in its vicinity and have more detailed knowledge of its ambient environment. CAVs can also form a platoon on the road, which can further increase road capacity, reduce energy consumption and improve traffic safety.

However, the transition from the current human-driven vehicles (HDVs) to a fully CAV environment may take a few decades. One report claims that all new vehicles will have the connectivity function by 2025 (16); another suggests that only 75% of vehicles will be autonomous by 2040 (17). The transition period will entail a “mixed” traffic flow of HDVs and CAVs. Different from the CAVs, human drivers in HDVs are full of personalization and uncertainty. Without the trajectory prediction, the CAV/CV will treat the ambient HDVs as a general type without the personalized consideration and only provide a short-time motion plan and control process, which may impair the advantage of CAV applications in a mixed flow context. Hence, HDV trajectory prediction is one of the most important challenges, as it can be leveraged to reduce traffic accident risks smooth traffic oscillations, and provide opportunities for applications like CAV platoon control, prediction-based autonomous emergency braking (AEB) system, and car-following guidance system.

For vehicle trajectory prediction, the input features can be extracted from the historical trajectory data of the target vehicle (the vehicle whose trajectory is being predicted) and its surrounding vehicles, e.g., the immediate leading vehicle. The output can be speed, acceleration rate or location (i.e. the longitudinal trajectory) of the target vehicle in a certain future time period. Prediction models include empirical equations with artificial parameters (18), extended
Kalman filtering models (19), feedforward neural networks (20), etc. Recurrent neural networks (RNNs), a state-of-the-art deep learning architecture, have been implemented in recent vehicle trajectory prediction studies (21–23). Deep learning models avoid manually-derived features and provide sufficient modeling complexity which guarantees promising performance when big data is available (24). RNNs are especially capable of modeling non-linear temporal dependencies in sequential data, which can mimic memory-based decision-making of human drivers. Long Short-Term Memory (LSTM) models are a variant of RNNs with the advantage of avoiding the vanishing gradient problem (25); they have been implemented in some vehicle trajectory prediction studies (26, 27).

Some limitations of previous studies make them unsuitable for the mixed HDV and CAV traffic flow scenario. First, most of them only rely on historical trajectory data of the target vehicle and its immediate leading vehicle. As shown in Figure 1(a), the previous speed of the target vehicle, and previous speed differences and spatial gaps between the two vehicles are extracted as model inputs (21–23, 26). One exception exists (27) that considers historical trajectories of nine surrounding vehicles for longitudinal velocity prediction. Some automotive manufacturers such as Tesla aim to predict future trajectories of other roads users merely based on board sensors of individual AVs (28), which may not be reliable in the complex real-world driving environment. As reported in (29), three Tesla drivers have died in crashes when the Autopilot system failed to detect and react to hazards. On the other hand, in a mixed traffic flow consisting of CAVs and HDVs as shown in Figure 1(b), the ego CAV (the CAV following the target HDV) can receive trajectory information from other CAVs over a much larger spatial range through V2V techniques; also, these CAVs possibly drive in different lanes. These new data sources need to be considered in the vehicle trajectory prediction model. Second, most previous studies simply make one-step predictions (21–23, 26). While this may be enough for applications like traffic simulation to replicate traffic patterns, a longer prediction horizon is more beneficial for CAV applications such as motion planning (30) and platoon control (31).

This study implements LSTM models to predict the target HDV’s longitudinal trajectories for multiple time steps for the scenario shown in Figure 1(b). Different from previous vehicle trajectory prediction studies (21–23, 26, 27), our LSTM models can utilize the historical trajectories from multiple vehicles, including the target HDV, and the ego CAV and CAVs within its communication range (e.g., $L = 200$ meters to the front and back in the left, middle and right lanes). This study further enhances the LSTMs using the attention mechanism, which was first proposed for RNNs on end-to-end machine translation applications (32). Previous studies have shown that the attention mechanism can improve deep learning model performance (33). Furthermore, it also increases the explanatory power of deep learning.

---

**Figure 1** (a) A pair of Target Vehicle and Leading Vehicle in previous studies (21–23, 26); (b) A mixed traffic flow of CAVs and HDVs considered in this study

---
models, which is as important as accuracy in modern artificial intelligence systems (34). In this study, the attention mechanism is used to combine historical trajectories of the target HDV and the relevant CAVs. The learned attention weights can indicate CAVs having larger impacts on the predictions, which help us to understand human driving behaviors and potentially reduce CAV data transmission burden (14, 15). The dataset used in this study is the Next Generation SIMulation (NGSIM) (3) data which has been widely used in previous studies (21–23, 26, 27).

The main contributions of this paper are:

- The attention-based LSTM models and two benchmark LSTMs are built and compared for multi-step HDV longitudinal trajectory prediction.
- The real-world NGSIM dataset is utilized to simulate a mixed HDV and CAV environment under various CAV market penetration rates (MPRs).
- The attention-based LSTM models make more accurate predictions through better utilization of historical trajectories from the target HDV, the ego CAV, and all CAVs within the ego CAV’s communication range.
- The grid-level average attention weights are analyzed to identify the CAVs that have larger impacts on the target HDV’s future trajectories.

The rest of the paper is organized as follows. Section 2 reviews related studies on vehicle trajectory prediction. Section 3 introduces the attention-based LSTM model. In Section 4, the processing of the NGSIM dataset is discussed. Section 5 illustrates the HDV trajectory prediction performance under various MPR scenarios. The paper concludes with a discussion on the study findings and future research directions.

LITERATURE REVIEW

There is substantial literature on understanding driving behaviors and predicting vehicle trajectories over the past few decades. This section introduces the relevant literature on two topics: microscopic traffic simulation and autonomous driving.

Microscopic Traffic Simulation

Microscopic traffic simulation is used to replicate driving behaviors and traffic phenomenon such as traffic breakdowns and stop-and-go oscillations (35). Car-following models and lane-changing models are key components in a microscopic traffic simulation model. The former focuses on capturing the longitudinal trajectories of a vehicle in response to its immediate leading vehicle and the latter seeks to produce reasonable lateral movement for the vehicle.

Most car-following and lane-changing models have forms of empirical equations with artificial parameters, such as the Gipps’ model (36) and IDM (18). Before the application, parameters in car-following and lane-changing models have to be calibrated to replicate real-world traffic conditions (37, 38). Although the form of these analytical models can guarantee fast simulation speed, their parsimony restricts model flexibility and accuracy. A detailed literature review of classical car-following models and lane-changing models can be found in Saifuzzaman and Zheng (2014) (39) and Rahman et al. (2013) (40), respectively.

Beyond the aforementioned analytical models, machine learning models like random forest models and neural network models have been proposed to model lane-changing in traffic simulation (41). Recently, RNN-based deep learning models have been applied to capture car-following behaviors to explain traffic flow phenomena like traffic oscillations (23) and asymmetric driving behaviors (21). These studies suggest that RNN models outperform classical IDM in terms of reproducing and explaining traffic flow patterns (21–23).

Autonomous Driving

Autonomous driving is another significant application domain for vehicle trajectory prediction. The development of all levels of autonomous driving systems such as Collision Warning System and Cooperative Adaptive Cruise Control System (13, 42) relies on the motion prediction of surrounding traffic participants. Vehicular technology companies such as Waymo, Uber, Tesla and Nvidia are all developing deep learning algorithms for AVs to predict future
trajecories of HDVs (28, 43–45). Compared to the replication of traffic patterns in traffic
simulation tools, autonomous driving has a much higher requirement in terms of vehicle
trajectory prediction accuracy.

Past studies can be split into two groups based on input data: physics-based motion
models and maneuver-based motion models (46). The former group makes trajectory
predictions purely based on the laws of physics. Acceleration rate, car weight, road surface
friction coefficient, etc., are usually taken into account. Examples include Constant Turn Rate
and Velocity (CTRV) and Constant Turn Rate and Acceleration (CTRA) models (47, 48). By
contrast, the latter group also considers the maneuver that the driver intends to perform. For
example, Houenou et al. (2013) propose a Maneuver Recognition Model (MRM) that can detect
the target vehicle’s maneuver such as keeping to the lane or making turns based on its
instantaneous path and the road shape. A Constant Yaw Rate and Acceleration motion model
(CYRA) is then applied to combine with the MRM to make final trajectory predictions (6).

The physics-based and maneuver-based models fail to model highly nonlinear terms
and can result in huge deviations from the real trajectory (7, 20). Recently, machine learning
models and deep learning models have been used to elicit more accurate predictions. Wiest et
al. (2012) build a Variational Gaussian Mixture Model to generate a whole distribution over
the future trajectories. An evaluation of the variance is used to examine the reliability of the
prediction (7). Yim and Oh (2004) propose a neural network model to predict lateral dynamics
of a vehicle such as the steering angle (20). In terms of deep learning models, LSTM models
have shown satisfying performance in multiple studies (26, 27, 49).

METHODOLOGY

Problem Description
In Figure 1(b), assume the ego CAV can communicate with $N$ CAVs within 200 meters front
and back in the left, current, and right lanes. Using CAV technologies, the historical $T$ time-
step trajectories of the $N$ CAVs, the ego CAV, and the target HDV are taken as the input to the
LSTM models. The task is to predict the future trajectories of the target HDV in the next $K$
time steps.

Attention-based LSTM for HDV longitudinal trajectory prediction
In this study, we propose an attention-based LSTM model to predict the target HDV trajectories
using historical trajectories of itself and the $(N+1)$ CAVs. Before discussing the proposed
LSTM model architecture, the basics of an LSTM model are briefly introduced hereafter.

A simple LSTM model for the target HDV trajectory prediction has an architecture as
shown in Figure 2. It consists of a sequence of LSTM cells. Each cell takes the corresponding
HDV trajectory data at that time step. The recurrence feature of the LSTM model allows it to
learn long-term relationships in sequential data. Hence, it can mimic decision-making processes
of human drivers and outperform analytical vehicle trajectory models (26, 27).
As shown in Figure 2(a), the recurrent architecture of the LSTM model can be unrolled into $T$ LSTM cells, the inputs to which are the trajectory data of the target HDV $[X_{t-T+1}^{HDV}, ..., X_t^{HDV}]$. At each time step $j$ ($j = t - T + 1, ..., t$), $X_j^{HDV}$ includes the $x$ and $y$ coordinates of the target HDV at time step $j$:

$$X_j^{HDV} = (x_j^{HDV}, y_j^{HDV})$$  \hspace{1cm} (1)

All of the LSTM cells share the same internal architecture. Figure 2(b) shows an example of a specific LSTM cell at time step $j$. Given the new input data at time step $j$ $X_j^{HDV}$, the cell state $C_{j-1}^{HDV}$ and the cell output $h_{j-1}^{HDV}$ from previous time step, the cell performs various operations to generate the new cell state $C_j^{HDV}$ and the cell output $h_j^{HDV}$. These operations are named as gates in the LSTM model. A forget gate determines how much information to “forget” from the cell’s previous cell state $C_{j-1}^{HDV}$:

$$f_j^{HDV} = \sigma(W_f \cdot [h_{j-1}^{HDV}, X_j^{HDV}] + b_f)$$  \hspace{1cm} (2)

An input gate decides the amount of new information to be stored in the memory based on $X_j^{HDV}$ and $h_{j-1}^{HDV}$:

$$i_j^{HDV} = \sigma(W_i \cdot [h_{j-1}^{HDV}, X_j^{HDV}] + b_i)$$  \hspace{1cm} (3)

Based on the forget and input gates, we can update the cell state:

$$C_j^{HDV} = \tanh (W_c \cdot [h_{j-1}^{HDV}, X_j^{HDV}] + b_c)$$  \hspace{1cm} (4)

$$C_j^{HDV} = f_j^{HDV} \cdot C_{j-1}^{HDV} + i_j^{HDV} \cdot C_j^{HDV}$$  \hspace{1cm} (5)
Finally, the output gate computes the new cell output based on the input $X_t^{HDV}, h_t^{HDV}$ and the updated cell state $c_t^{HDV}$:

$$o_t^{HDV} = \sigma(W_o \cdot [h_{t-1}^{HDV}, X_t^{HDV}] + b_o)$$  \hspace{1cm} (6)

$$h_t^{HDV} = o_t^{HDV} \cdot \tanh(c_t^{HDV})$$  \hspace{1cm} (7)

Note that $\sigma$ and $\tanh$ represent the sigmoid and hyperbolic tangent activation functions. The cell output $h_t^{HDV}$ from the last LSTM cell contains useful information from the whole $[X_{t-T+1}^{HDV}, ..., X_t^{HDV}]$ sequence, and is then fed as the input to a fully-connected feedforward layer, the output of which is the $s$ of the target HDV in the future $K$ time steps $[\tilde{X}_{t+1}^{HDV}, ..., \tilde{X}_{t+K}^{HDV}]$. $W_f, W_i, W_C, W_o, b_f, b_i, b_C, b_o$ are learned in the training process of the LSTM model.

Now assume that the historical trajectories of the CAV $v$ in the communication range of the ego CAV are $[X_{t-T+1}^{CAV}, ..., X_T^{CAV}]$, $v = 1, 2, ..., N$, and the ego CAV’s historical trajectories are $[X_{t-T+1}^{CAV}, ..., X_T^{CAV}]$. Based on the historical trajectories of the $(N + 1)$ CAVs and the target HDV, we design an attention-based LSTM model whose architecture is shown in Figure 3.

At time step $t$, the historical trajectories from the target HDV and the $(N + 1)$ CAVs are taken as inputs to the LSTM model, producing a hidden state matrix $H_t = [h_1^t, h_2^t, ..., h_N^t, h_t^{CAV}, h_t^{HDV}]$, $H_t \in \mathbb{R}^{d \times Q}$, where $d$ is the length of the hidden state in LSTM, and $Q = N + 2$ represents $(N + 1)$ CAVs and the target HDV. The attention pooling layer will first calculate an attention weight vector $\alpha_t = [\alpha_1^t, \alpha_2^t, ..., \alpha_N^t, \alpha_t^{CAV}, \alpha_t^{HDV}]$ based on $H_t$:

$$\alpha_t = \text{softmax}(\tanh(W_\alpha H_t))$$  \hspace{1cm} (8)

where $W_\alpha \in \mathbb{R}^{1 \times d}$ is also learned by the deep learning model.

The attention pooling layer then combines the historical trajectories from the CAVs and the target HDV through:

$$h_t = H_t (\alpha_t)^T$$  \hspace{1cm} (9)

$h_t \in \mathbb{R}^d$ is then utilized as the new input to a feedforward layer to calculate future trajectories $[\tilde{X}_{t+1}^{HDV}, ..., \tilde{X}_{t+K}^{HDV}]$.

**Figure 3** Architecture of the proposed attention-based LSTM model

Note that once the attention-based LSTM model is trained, the generated attention weight vector $\alpha_t = [\alpha_1^t, \alpha_2^t, ..., \alpha_N^t, \alpha_t^{CAV}, \alpha_t^{HDV}]$ can indicate the relative importance of the historical trajectories of the $(N + 1)$ CAVs and the target HDV. The larger the attention weight
of a vehicle is, the more important its historical trajectories are in terms of determining the future trajectories of the target HDV.

**DATASET**

The vehicle trajectory dataset used in this study is the publicly-available Next Generation Simulation (NGSIM) dataset collected at US-101. It includes 45 minutes of vehicle trajectories over a 0.3-mile freeway segment with five lanes. In NGSIM, GPS coordinates for each vehicle are recorded at 10 Hz. In this study, we use the vehicles’ local coordinates \((x, y)\) as the trajectory data, which are also provided in NGSIM. The origin is located at the entry point of the left-most edge of the section in the direction of travel. The \(y\)-axis is the longitudinal motion direction, and the \(x\)-axis is perpendicular to the \(y\)-axis. To reduce noise, we downsample the trajectory data by a factor of 2; so, each time step is 0.2-second long. We apply the attention-based LSTM models to predict the future 4-second trajectories \((K = 20 \text{ time steps})\) of the target HDV with the previous 3-second trajectory data \((T = 15 \text{ time steps})\) from the ego CAV, the target HDV, and the CAVs in communication range \((L = 200 \text{ meters})\).

The steps to process the raw NGSIM dataset are as follows:

1. Suppose there are \(N_{veh}\) vehicles in the US-101 dataset. Assume the CAV market penetration rate \(M_{PR_{CAV}}\) is one of four scenarios: 0.2, 0.4, 0.6 and 0.8. Label each vehicle as an HDV or a CAV based on the \(M_{PR_{CAV}}\).
2. Select a CAV following an HDV (the distance between them is less than \(L = 200 \text{ meters}\)), the reason of choosing 200 meters as the threshold is because this is the maximum detection range of automotive radar systems \((50, 51)\); the CAV is labeled the ego CAV, and the HDV is labeled the target HDV. Identify the CAVs that are in communication range of the ego CAV as shown in Figure 1(b).
3. At time step \(t\), the historical \(T\) time steps of trajectory data from the target HDV and the corresponding CAVs, as well as the \(K\) time steps of future trajectories of the target HDV form as one record in the dataset.
4. In total, 3,717,730 records are generated. The whole dataset is then split into training, validation and testing datasets in the ratio of \([70\%, 10\%, 20\%]\).

Note that the trajectory data is from HDVs in NGSIM. CAVs and HDVs may drive differently in real-world mixed traffic flow. However, such datasets are not available. Many studies are designing the driving behaviors of AVs by imitating human drivers based on HDV data \((32–34)\). Hence, the NGSIM-based data is a good proxy for mixed CAV and HDV trajectory data for this explorative study. Another approach is to apply traffic simulator-generated data according to calibrated driving behaviors of HDVs and specific CAV control strategies. However, this method will discard some properties of personalization. To consider as much information of individualized HDVs’ driving behavior as possible, the NGSIM dataset is applied since this study mainly focuses on the trajectory prediction of HDVs.

**EXPERIMENTAL RESULTS**

**HDV Trajectory Prediction Performance**

To analyze the performance of our attention-based LSTM models, two benchmark LSTM models are implemented. The first benchmark model is the naïve LSTM model that only uses historical trajectories from the ego CAV and the target HDV for future trajectory predictions. It is based on previous studies that only rely on trajectories from the target vehicle and its immediate leading vehicle \((21–23, 26)\). A comparison of the naïve LSTM and our attention-based LSTM models can verify whether the historical trajectories from surrounding CAVs can improve the trajectory prediction performance. The second benchmark model also utilizes the historical CAV trajectories. More specifically, the \([h_t^1, h_t^2, ..., h_t^N, h_t^{CAV}, h_t^{HDV}]\) hidden state vectors are concatenated as a large input vector to the feedforward layer instead of being fused through the attention pooling layer shown in Figure 3. This benchmark model is referred to as hidden-state-based LSTM (HS-LSTM). The optimal hyperparameters of these LSTM models are identified using the training and validation datasets through grid search. The length of the
hidden state $d$ in LSTM cells is set as 64. The neuron number of the feedforward layer is set as 128, and the sigmoid function is chosen as the activation function of this layer. The number of training epochs is 10 and the batch size is set as 128. All of the LSTM models are trained based on the Adam algorithm with a learning rate of 0.001. The data and implementation details for our experiments can be found at https://github.com/leilin-research/VTP.

Root Mean Square Error (RMSE) is used to evaluate these models. Assume that the target HDV’s local coordinate at time step $k$ is $x^H_{k}^{HDV} = (x^H_{k}^{HDV}, y^H_{k}^{HDV})$, $k = t + 1, ..., t + K$, and its trajectory prediction at the same time step is $\hat{x}^H_{k}^{HDV} = (\hat{x}^H_{k}^{HDV}, \hat{y}^H_{k}^{HDV})$. The displacement of the target HDV at time step $k$ can be calculated as:

$$
\hat{d}_k = \sqrt{(x^H_{k}^{HDV} - \hat{x}^H_{k}^{HDV})^2 + (y^H_{k}^{HDV} - \hat{y}^H_{k}^{HDV})^2}
$$

(10)

The RMSE can then be calculated as follows:

$$
RMSE = \frac{1}{K-M} \sum_{k=M}^{K} \sum_{m=1}^{M} (d_{km})^2
$$

(11)

where $K = 20$ is the prediction horizon, $M$ is the number of records in the testing dataset, and $d_{km}$ is the displacement between the prediction and the ground truth location of the vehicle at time step $k$ for the $m^{th}$ record.

Table 1 shows the RMSEs of the LSTM models for the testing dataset. The attention-based LSTM and the HS-LSTM are also tested for different CAV market penetration rates. First, it can be observed that the naïve LSTM has the worst performance, with the RMSE as high as 2.41. The trajectory prediction accuracy can be improved even when the $MPR_{cav}$ is only 0.2, e.g., the corresponding RMSEs for the HS-LSTM and the attention-based LSTM are 1.99 and 1.97, respectively. Second, as expected, the RMSE decreases when the $MPR_{cav}$ increases from 0.2 to 0.8 for both HS-LSTM and attention-based LSTM models. This indicates that more CAVs in the communication range can help the ego CAV to better capture the future driving behaviors of the target HDV. Finally, the attention-based LSTM has a slightly smaller RMSE compared to the HS-LSTM under all $MPR_{cav}$ scenarios. The statistical difference between the performance of HS-LSTM and attention-based LSTM is investigated using the Wilcoxon signed-rank test (55). The null hypothesis is that there is no statistical difference between the performance of the HS-LSTM and the attention-based LSTM at the significance level of $\alpha$. As can be observed, when the significance level $\alpha$ is 0.1, attention-based LSTM outperforms HS-LSTM for all $MPR_{cav}$ scenarios; when $\alpha$ is set as 0.01, attention-based LSTM still performs better than HS-LSTM when $MPR_{cav}$ is 0.2, 0.6, or 0.8. The learned attention weights can be further analyzed to identify CAVs that are more important for trajectory prediction, as discussed next.

<table>
<thead>
<tr>
<th>$MPR_{cav}$</th>
<th>Naïve LSTM</th>
<th>HS-LSTM</th>
<th>Attention-based LSTM</th>
<th>Wilcoxon signed-rank test ($p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>2.41</td>
<td>1.99</td>
<td>1.97</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>0.4</td>
<td>1.83</td>
<td>1.97</td>
<td>1.82</td>
<td>0.06*</td>
</tr>
<tr>
<td>0.6</td>
<td>1.76</td>
<td>1.71</td>
<td>1.61</td>
<td>&lt;0.001***</td>
</tr>
</tbody>
</table>

Significance level: *** $\alpha = 0.01$, * $\alpha = 0.1$

Figure 4 shows the RMSE for each of the 20 time steps for all three LSTM models. The RMSE is less than 1 meter for the first six prediction steps for all models. It keeps increasing with the time step. This is because the driving behaviors in the near future are more related to the historical data and are easier to predict compared to further in the future. For both HS-LSTM and attention-based LSTM, the model performance improves when the $MPR_{cav}$ increases from 0.2 to 0.8 as time step increases (e.g., from $t + 13$ to $t + 20$).
Figure 4 Comparison of RMSE per time step (meter) for three LSTM models

Grid-level Average Attention Weight Analysis

In attention-based deep learning models, the learned attention weights can be further analyzed to capture important features in the data (33). In this study, the normalized attention weight vector $\alpha_t$ in Equation (8) is used to combine the hidden states $H_t$ of the target HDV and CAVs. Hence, $\alpha_t$ can indicate which vehicles contribute more to the future trajectory of the target HDV for each prediction. To visualize the attention weights, we split the communication range of the ego CAV (e.g., 200 meters front and back in the left, current, and right lanes) into a 3 by 23 grid. We use “Left”, “Current”, and “Right” to represent the lateral positions of the grid cells and an index order from “-11” to “11” to represent the longitudinal positions. The grid cell where the ego CAV is located is labeled (Current, 0). Positive and negative longitudinal indices represent the grid cells in its front and back. The length of the front and back grid cells is about 18.28 meters (60 feet) long. When the $MPR_{CAV}$ is equal to 0.8, the average attention weight for each grid cell is calculated for the testing dataset. Figure 5 shows the results for scenarios in which the target HDV is located in the (Current, 1), (Current, 2), and (Current, 3) grid cells.
Figure 5 Spatial distribution of average attention weights by the location of the target HDV ($MPR_{CAV} = 0.8$): (a) the target HDV is in the (Current, 1) grid cell; (b) the target HDV is in the (Current, 2) grid cell; (c) the target HDV is in the (Current, 3) grid cell.

As can be seen in Figure 5, independent of the distance between the ego CAV and the target HDV, little impact is observed from CAVs located in the upstream (back) grid cells. The CAVs that are the most important to predict the target HDV’s future trajectory are those located in the same grid cell as the target HDV and those in the next downstream (front) grid cell in the same lane. For example, in Figure 5(c), the summation of the average attention weights from those two grid cells is 65.3% (51.2%+14.1%). For the remaining grid cells in the front of the target HDV, the average attention weight tends to decrease when the relative distance between the CAVs and the target HDV increases. This suggests that it may not be necessary to include all the information of the CAVs in communication range for the trajectory prediction. The proposed attention mechanism can potentially be useful for developing CAV communication protocols, e.g., reducing CAV data or prioritizing CAV data transmission (14, 15).

CONCLUDING COMMENTS

This study develops LSTM models to predict future longitudinal trajectories of a target human-driven vehicle (HDV) for an ego connected and autonomous vehicle (CAV) in a “mixed” traffic flow environment. A few LSTM variants including the Naïve LSTM, hidden-state-based LSTM (HS-LSTM), and the attention-based LSTM models are explored. The Naïve LSTM model considers only the historical trajectory data of the ego CAV and the target HDV, while the latter two models assume the ego CAV can utilize the CAVs that are in its communication range for trajectory prediction. The models are built and compared using the NGSIM dataset under different CAV market penetration rates ($MPR_{CAV}$). The results show that it is beneficial to include the information from CAVs within communication range in the LSTM models; the model performance can be improved even when the $MPR_{CAV}$ is only 0.2. The attention-based LSTM performs better than the HS-LSTM under all $MPR_{CAV}$ scenarios. Through grid-level analysis of the average attention weights, we determine that the CAVs in
the same grid cell as the target HDV and those in the next downstream (front) grid cell in the
same lane play a more significant role in predicting its trajectories.

Future research directions include testing our attention-based LSTM model on other
datasets such as highD, a drone-collected vehicle trajectory dataset (56). Also, the proposed
deep learning models can be extended for any available real-world mixed CAV and HDV
trajectory datasets through approaches such as transfer learning (57, 58). It is also interesting
to test our models based on mixed CAV and HDV data from traffic simulation models. Further,
it is important to evaluate the reliability of the LSTM models based on trajectory data with
various levels of noise or missing values. We also plan to capture uncertainties in trajectory
prediction using Bayesian deep learning techniques (59) and better combine trajectories from
CAVs with graph convolution neural networks (60). We will develop lane-changing prediction
models and conduct grid-level average attention weight analysis as well. The study insights are
also useful to develop control algorithms to optimize driving behaviors of a CAV platoon based
on HDV trajectory predictions (61).

AUTHOR CONTRIBUTIONS
The authors confirm contribution to the paper as follows: study conception and design: Lei Lin,
Siyuan Gong, and Srinivas Peeta; data processing: Lei Lin and Xia Wu; model development: Lei
Lin and Xia Wu; analysis and interpretation of results: Lei Lin; draft manuscript preparation: Lei
Lin, Siyuan Gong, and Srinivas Peeta. All authors reviewed the results and approved the final
version of the manuscript.

ACKNOWLEDGMENTS
This work is partially supported by National Key R&D Program of China (2019YFB1600100),
NSFC (71901038), Natural Science Basic Research Plan in Shaanxi Province (2020JQ-392), and
Research Funds for the Central Universities, Chang’an University (300102240301).

REFERENCES
2. Zhu, L., J. Gonder, and L. Lin. Prediction of Individual Social-Demographic Role Based on
Travel Behavior Variability Using Long-Term GPS Data. Journal of Advanced
2018.
4. Ossen, S., and S. Hoogendoorn. Car-Following Behavior Analysis from Microscopic
Trajectory Data. Transportation Research Record: Journal of the Transportation Research
Based Traffic Analysis: Methodological Framework for Automated Multivehicle Trajectory
Extraction. Transportation Research Record: Journal of the Transportation Research Board,
6. Houenou, A., P. Bonnifait, V. Cherfaoui, and W. Yao. Vehicle Trajectory Prediction Based on
Motion Model and Maneuver Recognition. Presented at the 2013 IEEE/RSJ International
Arizona, 2018.
Transportation Research Board 97th Annual Meeting Transportation Research Board, 2018.
10. SAE. J3016B: Taxonomy and Definitions for Terms Related to Driving Automation Systems
for On-Road Motor Vehicles - SAE International.
20. Frazzoli, E., M. A. Dahleh, and E. Feron. Robust Hybrid Control for Autonomous Vehicle...
Lin, Gong, Peeta, Wu


Lin, Gong, Peeta, Wu


