Traffic Forecasting using Vehicle-to-Vehicle Communication and Recurrent Neural Networks

Steven Wong
Department of Computer Science
University of California, San Diego
San Diego, CA 92110, USA
scw039@ucsd.edu

Robin Walters
Department of Computer Science
Northeastern University
Boston, MA 02115, USA
r.walters@northeastern.edu

Lejun Jiang, Tamas G. Molnar
Department of Mechanical Engineering,
University of Michigan, Ann Arbor
Ann Arbor, MI 48109, USA
{lejunj,molnart}@umich.edu

Rose Yu
Department of Computer Science
University of California, San Diego
San Diego, CA 92110, USA
roseyu@ucsd.edu

Abstract

We use vehicle-to-vehicle communication to make individualized traffic predictions on board of connected vehicles. Traditionally, traffic predictions rely on physics-based traffic models, where the propagation of congestion waves is simulated. Here we propose a hybrid model in which we supplement physics-based models with deep learning. We use a recurrent neural network to correct the errors of physics-based predictions and improve overall accuracy. Our model predicts velocity up to 40 seconds into the future with improved accuracy over physics-based baselines.

1 Introduction

The ability to predict future slowdowns on highways is beneficial both to individual vehicles and the highway population collectively. Vehicles may use such information to drive more smoothly improving safety, comfort, fuel economy and overall traffic throughput. Vehicle-to-vehicle (V2V) connectivity allows for individualized and more accurate predictions on board of connected vehicles as it enables a connected vehicle (which we call *ego*) to obtain information from vehicles ahead (called *lead*); see Fig. 1(a). The lead car's past data may then help predict the future of the ego car, as the ego will encounter the traffic the lead has already met. Such prediction can be done by physics-based models [1, 2, 5, 8, 9, 12, 16, 17, 18, 21], which capture the propagation of congestion waves along the highway, or by data-driven methods [3, 10, 14, 20]. Here we combine physics-based models with data-driven methods in a hybrid approach.

We use data collected by connected vehicles on public highways [15] to make individualized traffic predictions with a horizon of up to 40 seconds. Related work using deep neural networks makes either longer-term large-scale traffic predictions using road sensor data [10, 14], or very short-term predictions using video or LIDAR data [3, 20].

In this paper, we propose a hybrid method in which a recurrent neural network is trained to learn and correct the errors in the prediction of a physics-based model. Unlike a purely deep learning approach, our method leverages first principles from physics and insights from the study of traffic flow. At the same time machine learning may discover higher order correlations that are not captured by the first principle model, tune the model to the specific traffic conditions at hand, or improve robustness to

Machine Learning for Autonomous Driving Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

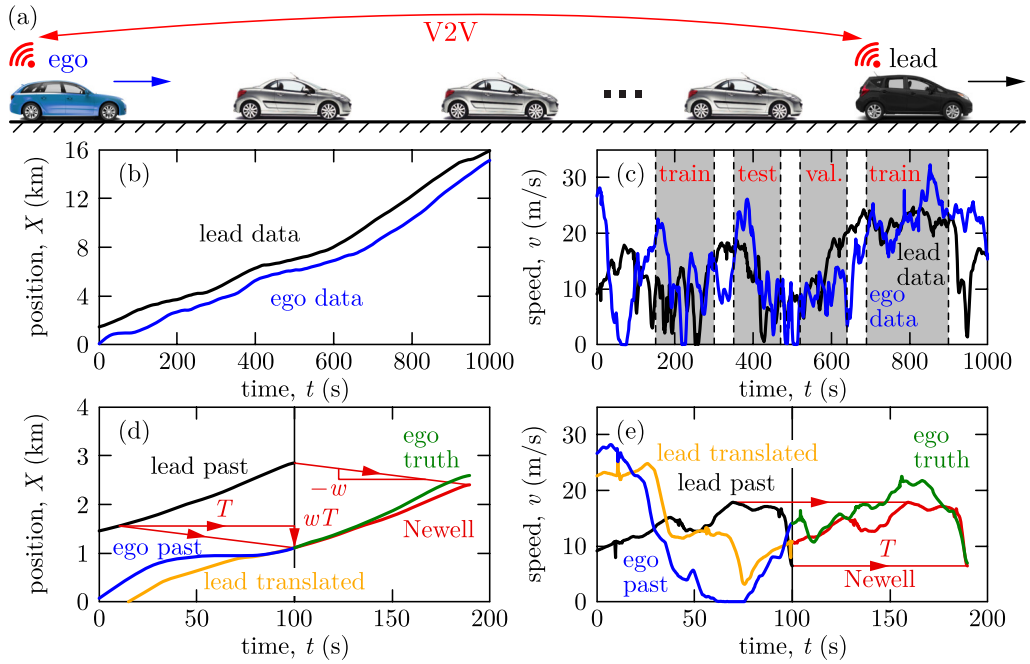

Figure 1: (a) Connected vehicles traveling on the highway where an ego vehicle intends to predict its future motion based on a lead vehicle's past data. (b,c) Experimental position and speed data. (d,e) The simplest first principle-based position and speed prediction.

unprocessed noisy signals. Our method achieves better accuracy than either a purely physics-based baseline or a pure machine learning approach using similar input features.

2 Background in Traffic Models

We focus on individualized traffic forecasting by predicting the trajectory of a connected vehicle (called ego) based on position and speed data obtained from another connected vehicle ahead (called lead) using V2V communication. Figure 1(b,c) show experimental position and speed data that were measured via GPS on connected vehicles driven on public roads as described in [15]. Since the ego vehicle undergoes qualitatively similar speed fluctuations as the lead, it allows the ego to predict its future motion based on the lead vehicle's past data transmitted through connectivity.

2.1 Physics-based Models

There exist various physics-based models to describe vehicular traffic. These include car-following models such as the OVM [2], the IDM [19] and models with time delays [8, 17], and traffic flow models like the LWR model [12, 18] and its novel formulations [9], the CTM [5] and the ARZ model [1, 21]. An elementary car-following model is the one by Newell [16], in which each vehicle copies the motion of its predecessor with a shift in time and space. Accordingly, the ego vehicle's position X_{E} is the same as the lead's position X_{L} with a shift T in time and a shift wT in space:

$$X_{\mathrm{E}}(t) = X_{\mathrm{L}}(t - T) - wT, \tag{1}$$

where w is the speed of congestion waves in traffic. Newell's model allows one to predict the ego vehicle's future motion as illustrated in Fig. 1(d,e). Using data about the lead and ego vehicles' past motion (black and blue) up to the prediction time t (vertical line), one can identify the time shift T between the trajectories by solving (1) for T. Then, the ego's future speed $v_{\mathrm{E}}(t+\theta)$ over a horizon $\theta \in [0, T]$ (green) is predicted by translating the lead vehicle's past trajectory (red):

$$\tilde{v}_{\mathrm{E}}(t, \theta) = v_{\mathrm{L}}(t + \theta - T). \tag{2}$$

Figure 1(d,e) illustrate that such a simple prediction (with $w = 5\,\mathrm{m/s}$) can qualitatively capture an upcoming slowdown, although there are quantitative errors in the ego's speed preview. One can

potentially improve prediction via more sophisticated first-principle models, however, the uncertainty of human driver behavior makes it challenging to achieve low prediction errors.

2.2 Related Work in Deep Learning

Using V2V connectivity to predict traffic by means of deep learning is a new area. Wang et al. use V2V communication for perception around obstacles and for making short-term trajectory predictions in urban environments [20]. Similarly, [3, 6, 11] make short-term predictions in the order of 3 seconds for individual car trajectories in urban traffic using sensor data and not V2V data. Our work focuses on making intermediate-term predictions (10-40 sec) of highway traffic via V2V connectivity.

Other works focus on long-term (10-20 min or longer) traffic prediction at certain road locations by considering large-scale road networks. Various architectures have been proposed in the literature to achieve this goal, such as 2D CNNs [14], recurrent convolutional architecture [10], graph convolutional networks [4], general regression neural networks [13] or spatio-temporal LSTM [22]. These predictions are usually lower in spatial and temporal resolution.

3 Methods

In what follows, we propose a hybrid method in which an a priori prediction from Newell's model is improved by deep learning. In this way, we maximally leverage the insights of the physics-based model and focus on learning higher order fluctuations in the signal. We explore two configurations of this method: (a) using the Newell prediction as additional input feature, called `VelLSTM-FC`, and (b) outputting the residual between the Newell prediction and the ground truth, called `ResLSTM-FC`.

We make the predictions in discrete time $t_i = i\Delta t$, $i \in \mathbb{Z}$ with a time step of $\Delta t = 0.1\,\mathrm{s}$. For `VelLSTM-FC`, the input $\mathbf{x}$ consists of two time series: k time steps of the ego's past velocity $v_\mathrm{E}(i-k+1), \dots, v_\mathrm{E}(i)$ and l time steps of the Newell prediction $\tilde{v}_\mathrm{E}(i,1), \dots, \tilde{v}_\mathrm{E}(i,l)$. The output $\mathbf{y}$ is l time steps of the ego's predicted future velocity $\hat{v}_\mathrm{E}(i,1), \dots, \hat{v}_\mathrm{E}(i,l)$. For `ResLSTM-FC`, the input $\mathbf{x}$ also includes the residual $R(i,j) = v_\mathrm{E}(i+j) - \tilde{v}_\mathrm{E}(i,j)$ of Newell's model $R(i,-k+1), \dots, R(i,0)$, while the output $\mathbf{y}$ is the next l time steps of the predicted residual $\hat{R}(i,1), \dots, \hat{R}(i,l)$. Then, the speed prediction $\hat{v}_\mathrm{E}$ is reconstructed as $\hat{v}_\mathrm{E}(i,j) = \hat{R}(i,j) + \tilde{v}_\mathrm{E}(i,j)$.

Model Architecture `VelLSTM-FC` and `ResLSTM-FC` use the same encoder-decoder architecture. We use a two-layer LSTM encoder with `ReLU` and $\tanh$ activation functions [7] to encode input into a hidden state vector $h = \texttt{LSTM-Enc}(\mathbf{x})$ where we initialize the hidden state vector randomly. The input vectors are concatenated into a single sequence: $\mathbf{x} = [v_E(i-k+1:i)\ \ \tilde{v}_E(i,1:l)]$ for `VelLSTM-FC` and $\mathbf{x} = [v_E(i-k+1:i)\ \ \tilde{v}_E(i,-\underline{k}+1:l)\ \ R(i,-k+1:0)]$ for `ResLSTM-FC`. This structure accommodates different numbers of time steps for different inputs. The hidden state vector h is then passed to a linear decoder layer which predicts l future time steps in one shot $\mathbf{y} = \texttt{Dec}(h)$. We found that predicting multiple time steps at once helps relative to predicting one time step autoregressively. The associated hyperparameters were tuned in the ranges listed in Appendix A.

4 Results

Dataset Data were collected by driving five connected vehicles in the Wednesday morning peak-hour traffic along US39 near Detroit, Michigan [15]. Each vehicle measured its position and speed by GPS, which were sampled and transmitted amongst the vehicles every 0.1s. Two vehicles travelled farther ahead of the other three with average distance of around 1300 and 900 m. This separation created six lead-ego vehicle pairs, and we use 1000 s data from two distant pairs under various traffic conditions. We constructed samples using rolling windows with 600 input and 400 output frames. Train, validation, and test data were separated as in Fig. 1(c) and normalized based on the train set.

Prediction Accuracy We compute two performance metrics: the velocity error (VE) at a specific prediction horizon $\theta_j = j\Delta t$ and the average velocity error (AVE) that is a mean over various horizons, for a given set $\mathcal{I}$ of samples (e.g. the test set) with $|\mathcal{I}|$ elements:

$$\mathrm{VE}(j) = \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} |\hat{v}_\mathrm{E}(i,j) - v_\mathrm{E}(i+j)|, \qquad \mathrm{AVE}(j) = \frac{1}{j}\sum_{m=1}^{j} \mathrm{VE}(m). \tag{3}$$

Method	VE@10s	VE@20s	VE@30s	VE@40s	AVE@40s
Constant Velocity	**3.94**	5.10	5.69	6.78	4.61
Newell (translated)	5.37	5.40	5.08	**4.77**	5.24
`Vel-FC`	4.59 ± 0.38	5.88 ± 0.27	6.36 ± 0.48	7.38 ± 0.49	5.48 ± 0.15
`Ego-only LSTM-FC`	5.45 ± 0.44	7.33 ± 0.63	7.86 ± 0.86	7.03 ± 0.81	6.23 ± 0.48
`VelLSTM-FC`	4.67 ± 0.86	6.15 ± 0.54	6.63 ± 0.54	6.63 ± 0.88	5.43 ± 0.75
`ResLSTM-FC`	4.24 ± 0.31	$\mathbf{4.49} \pm 0.18$	$\mathbf{4.98} \pm 0.26$	5.79 ± 0.57	$\mathbf{4.34} \pm 0.14$

Table 1: Results for the test set ($t \in [350, 470]$ seconds) shown in Fig. 1(c). Forecasting accuracy at $\theta = 10, 20, 30, 40$ seconds and mean accuracy over 40 seconds is shown. All numbers are in m/s. Errors are shown as mean plus minus one standard deviation over five trainings with random seeds.

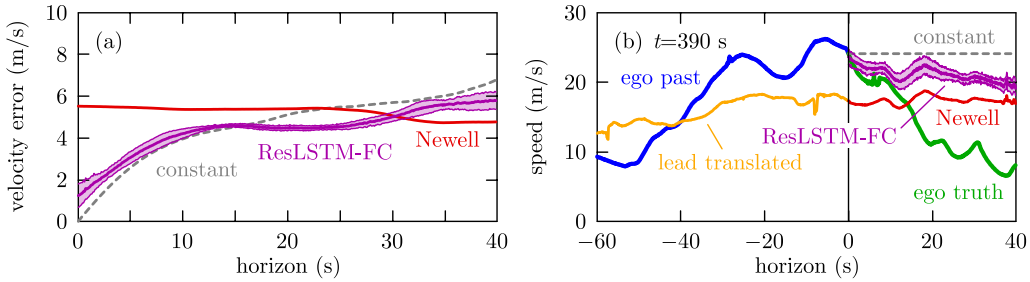

Figure 2: (a) Velocity error versus prediction horizon for constant speed preview (grey), Newell's model (red) and `ResLSTM-FC` method (purple). (b) Predictions compared to the ground truth (green) for a selected sample. Purple shading shows the standard deviation of results over five trained models.

Based on these metrics, we compare the performance of the proposed method to the following baselines: constant speed prediction, Newell's model, and pure deep learning in which only the ego's velocity is used as input (called `Ego-only LSTM-FC`). Table 1 and Fig. 2(a) summarize our results, while a sample prediction is illustrated in Fig. 2(b).

The prediction accuracy decreases over the prediction horizon with different rates for different methods. The constant speed prediction is accurate over short horizons, but has rapid error growth, while Newell's error starts higher and grows less. The `ResLSTM-FC` method combines the advantages of both of these baselines and outperforms them for prediction horizons of $15 \leq \theta \leq 30$ seconds. The `Ego-only LSTM-FC` method performs significantly worse than `ResLSTM-FC`, which validates our hybrid approach and shows the benefits of data from V2V connectivity. Lastly, we compare `ResLSTM-FC` to a monolithic fully connected network (called `Vel-FC`; see Appendix B) which has the same input and output as `VelLSTM-FC`. These networks have similar performance, and they are both outperformed by `ResLSTM-FC`, which gives the overall best average velocity error. This shows that, apart from supplementing deep learning with the physics-based Newell's model, predicting the residual is also an important element of our method.

5 Discussion

Hybrid models that integrate physics-based and data-driven elements were constructed for traffic prediction utilizing vehicle-to-vehicle communication. The model `ResLSTM-FC` showed improved accuracy over both purely first principle-based and purely deep learning-based baselines. This model used Newell's prediction as input to an LSTM neural network and the prediction of the residual error as output. `ResLSTM-FC` shares the best characteristics of each of our numerical baselines, having small error for short-term predictions and having slow growing error as forecasting horizon increases. The model was trained and tested on unprocessed GPS data, and the observed robustness makes it a feasible candidate for real-time on-board traffic predictions for connected vehicles.

We hypothesize that further improvements in our model can be obtained by more careful corrections of errors in the underlying physics model. Namely, while the Newell prediction often succeeds in predicting the degree of a slowdown, it is often slightly early or late. This corresponds to a time-shift

error. Future work includes adding prediction of this time shift or time warp to `ResLSTM-FC`. Another potential area for improvement would be to replace the Newell model with a more sophisticated first-principle baseline.

Acknowledgments

This research was partially supported by the University of Michigan's Center of Connected and Automated Transportation through the US DOT grant 69A3551747105, Google Faculty Research Award, NSF Grant #185034, and the U. S. Army Research Office under Grant W911NF-20-1-0334.

References

[1] A. Aw and M. Rascle. Resurrection of "second order" models of traffic flow. *SIAM Journal on Applied Mathematics*, 60(3):916–938, 2000.

[2] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, 51(2):1035–1042, 1995.

[3] M.-F. Chang, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, and J. Hays. Argoverse: 3D tracking and forecasting with rich maps. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8748–8757, 2019.

[4] Z. Cui, K. Henrickson, R. Ke, and Y. Wang. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 21(11):4883–4894, 2020.

[5] C. F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4):269–287, 1994.

[6] J. Gao, C. Sun, H. Zhao, Y. Shen, D. Anguelov, C. Li, and C. Schmid. Vectornet: Encoding HD maps and agent dynamics from vectorized representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11525–11533, 2020.

[7] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[8] Y. Igarashi, K. Itoh, K. Nakanishi, K. Ogura, and K. Yokokawa. Bifurcation phenomena in the optimal velocity model for traffic flow. *Physical Review E*, 64(4):047102, 2001.

[9] J. A. Laval and L. Leclercq. The Hamilton-Jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B: Methodological*, 52:17–30, 2013.

[10] Y. Li, R. Yu, C. Shahabi, and Y. Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In *International Conference on Learning Representations (ICLR)*, 2018.

[11] M. Liang, B. Yang, R. Hu, Y. Chen, R. Liao, S. Feng, and R. Urtasun. Learning lane graph representations for motion forecasting. *arXiv preprint arXiv:2007.13732*, 2020.

[12] M. J. Lighthill and G. B. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A - Mathematical, Physical and Engineering Sciences*, 229(1178):317–345, 1955.

[13] I. Loumiotis, K. Demestichas, E. Adamopoulou, P. Kosmides, V. Asthenopoulos, and E. Sykas. Road traffic prediction using artificial neural networks. In *Proceedings of the South-Eastern European Design Automation, Computer Engineering, Computer Networks and Society Media Conference*, pages 1–5, 2018.

[14] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. *Sensors*, 17(4):818, 2017.

[15] T. G. Molnár, D. Upadhyay, M. Hopka, M. Van Nieuwstadt, and G. Orosz. Delayed Lagrangian continuum traffic models for on-board traffic prediction. *Transportation Research Part C: Emerging Technologies*, 2020. Under review.

[16] G. F. Newell. A simplified car-following theory: a lower order model. *Transportation Research Part B: Methodological*, 36(3):195–205, 2002.

[17] G. Orosz, R. E. Wilson, and G. Stépán. Traffic jams: dynamics and control. *Philosophical Transactions of the Royal Society A*, 368(1928):4455–4479, 2010.

[18] P. I. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956.

[19] M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2):1805, 2000.

[20] T.-H. Wang, S. Manivasagam, M. Liang, B. Yang, W. Zeng, and R. Urtasun. V2VNet: Vehicle-to-vehicle communication for joint perception and prediction. In *European Conference on Computer Vision*, pages 605–621. Springer, 2020.

[21] H. M. Zhang. A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, 36(3):275–290, 2002.

[22] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2):68–75, 2017.

A Hyperparameter Tuning

We tuned hyperparameters over the ranges shown in Table 2.

Hyperparameter	Range
Input Time Steps (k)	300, 400, 500, **600**
LSTM Hidden Units	10, 20, 50, 100, **200**
Learning Rate	**0.0005**, 0.001, 0.005, 0.01
Prediction Horizon (l)	50, 100, 200, 300, **400**

Table 2: Hyperparameter tuning ranges. Bold values correspond to best performance.

B Vel-FC Model Architecture

The model architecture of `Vel-FC` is shown in Table 3.

Layer	Hyperparameters
Linear 1	in_features=1000, out_features=200, bias=True, followed by ReLU
Linear 2	in_features=200, out_features=200, bias=True, followed by ReLU
Linear 3	in_features=200, out_features=400, bias=True

Table 3: Model architecture of `Vel-FC`.