Initial Scene Configurations for Highway Traffic Propagation

Tim A. Wheeler
Aeronautics and Astronautics
Stanford University
Stanford, California 94305
wheelert@stanford.edu

Philipp Robbel
Robert Bosch LLC
Palo Alto, California 94304
philipp.robbel@us.bosch.com

Mykel J. Kochenderfer
Aeronautics and Astronautics
Stanford University
Stanford, California 94305
mykel@stanford.edu

***Abstract*—Validation of automotive safety systems can be done by simulating millions of driving traces. It is important that the distribution of initial scenes for these driving traces be as representative of reality as possible so that safety risk can be estimated accurately. This paper presents a methodology for constructing probability distributions over initial highway scenes from which samples can be drawn for safety evaluation through simulation. A method for automated model construction based on a Bayesian statistical framework is introduced and applied to the NGSIM Highway 101 and Interstate 80 datasets. Four models of increasing complexity and fidelity are developed. A complete implementation is available online.**

I. Introduction

As the automotive industry moves towards autonomous driving, it becomes increasingly necessary to develop advanced driving safety systems, collision prediction systems, and the tools for their rigorous analysis. Regulatory authorities, such as the European Commission and the National Highway Traffic Safety Administration, require drive tests to ensure system effectiveness and safety. Although a drive test can evaluate a system in actual operation, only relatively few cases can be examined due to time, cost, and safety constraints. Simulation can be used to assess system robustness over a significantly wider range of situations.

Recently, models of aircraft behavior have been developed to validate aircraft collision avoidance systems. These models used dynamic Bayesian networks to represent encounters between aircraft [1]. Both the model structure and parameters were learned from a large corpus of radar data. These models have been critical in the development and validation of a next-generation collision avoidance system [2]. This paper adapts this approach from aviation and uses it in the derivation of probabilistic driving models.

Probability theory has long been used in traffic flow research to capture statistics for macroscopic flow behavior and microscopic behavior models. Headway spacing and car-following models have been studied extensively since the 1950s [3], [4], with more recent models leveraging interactions among a large number of vehicles to obtain dynamic distributions over traffic fluctuations [5]. These models typically assume vehicles perfectly follow their lanes and thus do not capture the complete position and orientation needed for proper validation of a safety system using high-fidelity propagation models.

Modern traffic simulation programs vary in how they initialize traffic scenes. Simulation of Urban Mobility (SUMO), an open-source microscopic traffic simulator developed by the German Aerospace Center (DLR), initializes vehicles using univariate Gaussian speed distributions and only inserts vehicles if minimum headway requirements are met. Vehicles adhere to the lane centers and lane changes are instantaneous [6]. MITSIM, an open-source microscopic traffic simulator developed at the Massachusetts Institute of Technology Intelligent Transportation Systems Program, also inserts vehicles based on a minimum headway but sets the initial speed deterministically based on mean section speed and the vehicle's desired speed [7]. More advanced proprietary systems, such as PreScan by Tass International, provide tools for the automatic generation of origins and destinations but use similar spawning strategies to populate road networks.

The probabilistic scene distributions developed in this paper allow for the stochastic sampling of initial scenes for a straight section of highway that accurately represents the distributions of scenes learned from real-world data. The resulting models describe complete vehicle poses including centerline offsets and orientations. The proposed method can be used to generate scenes across varying lane counts and the resulting models can be sampled as many times as needed to obtain initial scenes for safety system validation through simulation. The performance of the developed methods are evaluated on two real-world datasets.

II. Model Overview

Automotive safety system analysis through simulation requires the stochastic sampling of initial driving scenes. Let a scene s define a joint configuration of traffic participants. A scene distribution $P(s \mid h)$ is a probability distribution over scene configurations given a highway topology h. This work seeks to establish a framework for obtaining an accurate representation of $P(s \mid h)$ from real-world driving data.

It is desirable that a model for representing scene distributions over a highway section for use in Monte Carlo safety system analysis possess several properties. First, the model should adequately capture the correlations between a variable number of vehicles in the continuous, multidimensional state space defined by the given highway section. Second, the model

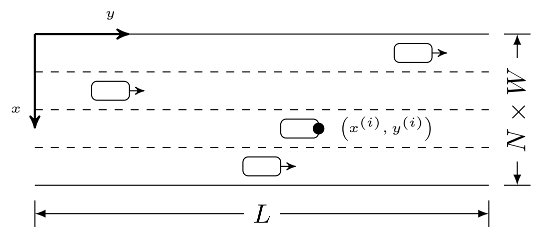


Fig. 1: Lane-relative frame for a straight multi-lane highway

should allow for the efficient sampling of scenes. Sampled scenes should be distributed as closely as possible to the true distribution the model is meant to represent. Finally, the model should be capable of computing scene likelihoods so that it can be evaluated against other models using cross-validation.

Directly modeling the scene distribution is made difficult by varying traffic participant counts, all of which possess varying degrees of correlation. The method adopted in this work is to only model correlations between vehicles in a local neighborhood. We show that such a model can be used to iteratively construct a scene by inserting vehicles along the fringes until the entire highway section is populated. The resulting method can be applied to highway sections with arbitrary lane counts and section lengths.

The models developed in this paper are constructed using several assumptions. First, all models assume a straight section of multi-lane highway. Second, all models assume that lanes can be freely shifted longitudinally with respect to one another. Third, the models do not distinguish between vehicle classes such as trucks and motorcyclists. Relaxation of these assumptions is left for future work.

Consider a straight highway section h of length L and N lanes of width W. A scene s is an arrangement of m vehicles $\vartheta^{(i)}$, $i = 1, \ldots, m$ within the bounds of the highway section. Vehicles are described by four-tuples $\vartheta^{(i)} = \langle x, y, v, \phi \rangle$ containing the two position coordinates, speeds, and lane-relative headings. The x-axis measures the distance of the center leading point of the vehicle to the left-most edge of the highway and the y-axis measures the distance of the same point from the scene entry point, as shown in Fig. 1.

This work uses Bayesian networks to represent probability distributions. Bayesian networks leverage conditional independence between variables to reduce the number of parameters required to define a joint distribution. Statistical techniques in the machine learning community readily infer model structure from data, and efficient algorithms exist for computing and sampling from marginal and conditional distributions for Bayesian networks [8].

Many off-the-shelf structure learning algorithms for Bayesian networks require discrete variables. Continuous and hybrid variables in this work are discretized and modeled using multinomial distributions. The distributions over values within each bin are assumed to be uniform. This approach allows models to match arbitrarily complex distributions with a sufficient number of bins, and has been successfully applied in aircraft encounter models [1].

Four models are developed in this work; a *marginal* model defining marginal probabilities over vehicle state variables, a *base* model representing a joint distribution over the vehicle state variables, a *chain* model with additional distance and speed variables to capture correlations between adjacent vehicles within lanes, and a *hierarchical* model which relates lane distributions using a global scene density and lane positions. Examples of randomly selected real-world and sampled scenes from the hierarchical model are shown in Fig. 2.

A. Marginal Model

The marginal model defines marginal distributions over a vehicle's velocity $p(v)$, centerline offset $p(d_{cl})$, lane-relative heading $p(\phi)$, and headway distance $p(d_{\text{front}})$, and is represented as an edgeless Bayesian network. It is a baseline against which more complicated models can be measured. Lanes are assumed independent and are constructed individually:

For each lane $l \in 1, \ldots, N$:

1) Sample the first vehicle according to $\vartheta_x \leftarrow \text{center}(l) + d_{cl} \sim p(d_{cl})$, $\vartheta_v \sim p(v)$, and $\vartheta_\phi \sim p(\phi)$. Sample the distance to the next vehicle, $d_{\text{front}} \sim p(d_{\text{front}})$.
2) The longitudinal position of the first vehicle is shifted randomly based on the headway distance to a potential trailing vehicle, $\vartheta_y \leftarrow y_{\text{offset}} \sim U\left(0, d \sim p(d_{\text{front}})\right)$.
3) Construct the next vehicle at $y^{(i)} \leftarrow y^{(i-1)} + d_{\text{front}}$ by sampling from all four marginal distributions as long as the location of the next vehicle given d_{front} is within the scene section length L.

The joint probability over vehicles in a scene under a marginal model $\mathcal{M}_m$ factors over lanes. The likelihood of a single lane with n vehicles indexed in increasing longitudinal order is given by:

$$p(l \mid \mathcal{M}_m) = \left(\prod_{i=1}^{n} p(\phi^{(i)}) p(v^{(i)}) p(d_{cl}^{(i)})\right) \left(\prod_{i=1}^{n-1} p(d_{\text{front}}^{(i)})\right) \times p(y_{\text{offset}}) \cdot p(d_{\text{front}} > L - x^{(n)}) \tag{1}$$

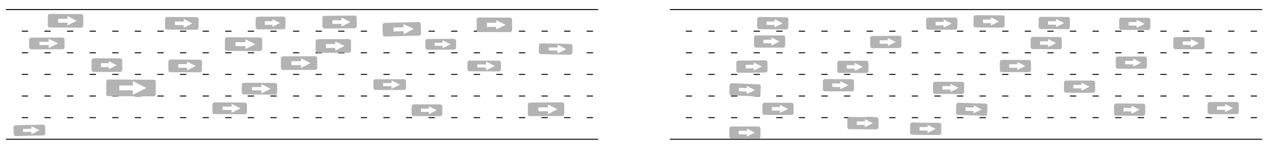
B. Base Model

The base model is a Bayesian network over the variables in the marginal model. It can capture correlations between the model variables.

For each lane $l \in 1, \ldots, N$:

1) Sample from the joint distribution $p(v, \phi, d_{cl}, d_{\text{front}})$.
2) Construct the first vehicle with $\vartheta_x \leftarrow \text{center}(l) + d_{cl}$, $\vartheta_v \leftarrow v$, and $\vartheta_\phi \leftarrow \phi$.
3) The longitudinal position of the first vehicle is shifted according to $\vartheta_y \leftarrow y_{\text{offset}} \sim U\left(0, d \sim p(d_{\text{front}})\right)$.
4) As long as the location of the next vehicle given d_{front} is within the scene section length L, sample from $p(v, \phi, d_{cl}, d_{\text{front}})$ and construct the next vehicle.

The joint probability over vehicles in a scene under base base model $\mathcal{M}_b$ factors over lanes. The likelihood of a lane with n vehicles indexed in increasing longitudinal order is:



(a) Real World

(b) Sampled from Hierarchical Model

Fig. 2: Example of real world and generated scenes from Highway 101, scene section 101b (defined in Section III)

$$p(l \mid \mathcal{M}_b) = \left(\prod_{i=1}^{n-1} p(v^{(i)}, \phi^{(i)}, d_{cl}^{(i)}, d_{\text{front}}^{(i)})\right) \cdot p(v^{(n)}, \phi^{(n)}, d_{cl}^{(n)}) \\ \times p(y_{\text{offset}}) \cdot p(d_{\text{front}} > L - x^{(n)} \mid v^{(n)}, \phi^{(n)}, d_{cl}^{(n)}) \tag{2}$$

C. Chain Model

The chain model extends the base model with three variables to further capture relations between successively generated vehicles. The additional variables are the distance to the trailing vehicle d_{rear}, the velocity of the trailing vehicle v_{rear}, and the velocity of the leading vehicle v_{front}.

For each lane $l \in 1, \ldots, N$:

1) Sample from the distribution marginalized by v_{rear}, $p(v, \phi, d_{cl}, d_{\text{front}}, v_{\text{front}}, d_{\text{rear}})$.
2) Construct the first vehicle with $\vartheta_x \leftarrow \text{center}(l) + d_{cl}$, $\vartheta_v \leftarrow v$, and $\vartheta_\phi \leftarrow \phi$.
3) The longitudinal position of the first vehicle is shifted according to $\vartheta_y \leftarrow y_{\text{offset}} \sim U(0, d_{\text{rear}})$.
4) As long as the location of the next vehicle given d_{front} is within the scene section length L, construct the next vehicle by sampling from $p(\phi, d_{cl}, d_{\text{front}} \mid d_{\text{rear}}, v_{\text{rear}} = v, v = v_{\text{front}})$ by conditioning on known values.

The joint probability over vehicles in a scene under a chain model $\mathcal{M}_c$ factors over lanes. The likelihood of a lane with n vehicles indexed in increasing longitudinal order is:

$$\begin{aligned} p(l \mid \mathcal{M}_c) = {} & p(v^{(1)}, \phi^{(1)}, d_{cl}^{(1)}, d_{\text{front}}^{(1)}, v_{\text{front}}^{(1)}) \\ & \times \left(\prod_{i=2}^{n-1} p(\phi^{(i)}, d_{cl}^{(i)}, d_{\text{front}}^{(i)}, v_{\text{front}}^{(i)} \mid d_{\text{rear}}, v, v_{\text{rear}})\right) \\ & \times p(\phi^{(n)}, d_{cl}^{(n)}, (d_{\text{front}}^{(n)} > L - x^{(n)}) \mid d_{\text{rear}}, v, v_{\text{rear}}) \\ & \times p(y_{\text{offset}} \mid v^{(1)}, \phi^{(1)}, d_{cl}^{(1)}, d_{\text{front}}^{(1)}, v_{\text{front}}^{(1)}) \end{aligned} \tag{3}$$

D. Hierarchical Model

The previous three models assume that lanes are generated independently, and thus can produce scenes with widely varying lane flow rates and densities. The hierarchical model attempts to address this by introducing ρ_{scene}, the number of vehicles per scene. Variation in traffic across lanes is captured using lane class C_{lane}, which indicates whether a given lane is the leftmost, rightmost, or bordered on both sides. The lane generation procedure is identical to that of the chain model, except that the scene density ρ_{scene} is sampled at the beginning of the generation procedure and is used with the lane class C_{lane} to condition on each distribution.

Computing the likelihood for a scene under a hierarchical model $\mathcal{M}_h$ requires integrating out the unobserved ρ_{scene},

$$p(s \mid \mathcal{M}_h) = \int_{\rho_{\text{scene}}} p(\rho_{\text{scene}}) \left[\prod_l p(l \mid \rho_{\text{scene}}, C_{\text{lane}}^{(l)})\right] d\rho_{\text{scene}}, \tag{4}$$

where the likelihood of a lane is computed as for the chain model but conditioned on the scene density and lane class.

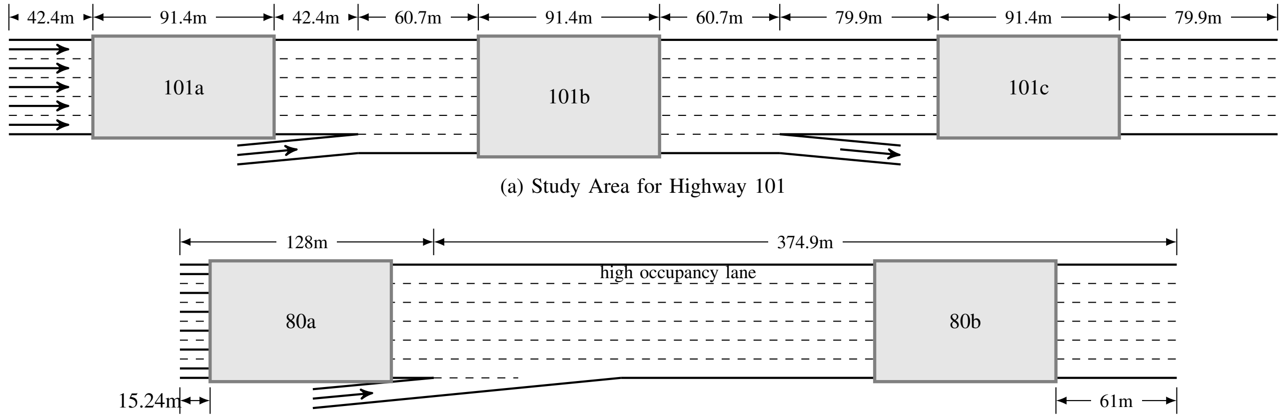
III. Data Source

This work used real-world driving data obtained from the Next-Generation Simulation (NGSIM) US Highway 101 and Interstate 80 datasets [9], [10]. Each dataset consists of 45 minutes of vehicle trajectory data collected using synchronized digital video cameras providing the vehicle lane positions and velocities over time at 10Hz. The US Highway 101 dataset covers an area in Los Angeles, CA, approximately 640m in length with five mainline lanes and a sixth auxiliary lane providing highway entrance and exit. The Interstate 80 dataset covers an area in the San Francisco Bay Area approximately 500m in length with six mainline lanes, including a high-occupancy vehicle lane and an onramp. These datasets were collected by the Next-Generation Simulation program in 2005 to facilitate automotive research and are freely available.

Traffic density in the datasets transitions from uncongested to full congestion and exhibits a high degree of vehicle interaction as vehicles merge on and off the highway and must navigate in the nearly-congested flow. This and the datasets' complete scene description make these sources particularly useful for learning traffic scene distributions.

The NGSIM datasets provide positions and velocities in the lane-relative frame as shown in Fig. 1. Vehicle trajectories were smoothed according to the method described by Thiemann *et al.* [11]. Vehicle centerline offset d_{cl} is defined as the signed lateral offset of a vehicle from the closest lane centerline. Lane-relative headings ϕ are obtained by assuming zero sideslip and estimating the heading directly from the trajectory, $\phi = \tan^{-1}(\Delta y / \Delta x)$. Headway distances were measured from the front of the trailing vehicle to the rear of the leading vehicle. All vehicles generated by the models used the average vehicle length and width of 4.34 and 2.06 meters respectively.

Scenes were extracted by pulling all vehicles within a given highway section from the dataset. In the Highway 101 dataset,



(a) Study Area for Highway 101

(b) Study Area for Interstate 80

Fig. 3: Data source schematics with scene sections. Each scene section is 91.4m long.

three highway sections were considered: section 101a before the auxiliary lane, section 101b including the auxiliary lane, and section 101c after the auxiliary lane. For Interstate 80 two highway sections were considered: section 80a before the onramp and section 80b after the onramp. These sections are overlaid in Fig. 3. All scenes are 91.4m (300 ft) long. Scenes were subsampled at 0.1Hz to decrease sample correlation.

IV. Model Learning

The models considered in this paper are parameterized by a set of hyperparameters λ and model parameters θ. Given a candidate model $\mathcal{M}(\lambda, \theta)$, a hyperparameter instantiation, and a training dataset $\mathcal{D}$ of m traffic scenes $s_{1:m}$, the model parameters are tuned to maximize the likelihood

$$\theta^* = \arg\max_\theta P(\theta \mid \mathcal{D}, \mathcal{M}, \lambda). \tag{5}$$

An application of Bayes' rule results in

$$P(\theta \mid \mathcal{D}, \mathcal{M}, \lambda) = \frac{P(\mathcal{D} \mid \theta, \mathcal{M}, \lambda) P(\theta \mid \mathcal{M}, \lambda)}{P(\mathcal{D} \mid \mathcal{M}, \lambda)}, \tag{6}$$

where $P(\theta \mid \mathcal{M}, \lambda)$ is a prior distribution over model parameters and $P(\mathcal{D} \mid \mathcal{M}, \lambda)$ is constant in θ. If there is sufficient mixing time between samples such that scenes can be assumed independent and identically distributed, then the likelihood of the dataset is the product of the scene likelihoods,

$$P(\mathcal{D} \mid \theta, \mathcal{M}, \lambda) = \prod_{s_i \in \mathcal{D}} P(s_i \mid \theta, \mathcal{M}, \lambda) \tag{7}$$

If we furthermore assume a uniform prior over model parameters, the maximum a-posteriori instantiation of the parameter vector is given by

$$\theta^* = \arg\max_\theta P(\mathcal{D} \mid \theta, \mathcal{M}, \lambda) \tag{8}$$

The goal of this work is to obtain a model that best reflects the true distribution over highway scenes. A model can be evaluated on a given dataset according to its likelihood, $P(\mathcal{D} \mid \theta, \mathcal{M}, \lambda)$. The higher the probability of the observed scenes, the better the model is said to predict the data.

Each model depends on the model parameters θ estimated from the data and a set of hyperparameters λ. Assigning the hyperparameters to maximize the likelihood of the model given the training data often leads to overfitting. One method for preventing overfitting is to separate the data into training and validation sets. Model parameters for a candidate hyperparameter set were learned from the training set and then compared to other hyperparameter instantiations by their likelihood score on the disjoint validation set.

Ten rounds of 10-fold cross-validation were conducted to reduce the variance caused by training on random partitions of the dataset. The dataset was randomly divided into 10 equal-sized sets, a model was trained against 9 of those parts and validated against the remaining part for each of the 10 possible allocations, and the average cross-validated likelihood across all 10 parts was computed. The average cross-validated likelihood across rounds is distributed according to a Student's-t distribution [12]. Scene distribution models were compared using their average cross-validated likelihoods. In practice one uses the log likelihood to avoid numerical precision problems.

The model parameters for a Bayesian network are the network structure and the conditional probability table counts for each variable. The problem of finding the graph structure of a Bayesian network that maximizes the likelihood is NP-complete [13], but efficient heuristic methods exist in the machine learning literature to find approximately optimal solutions [14], [15]. Once a model structure is selected, the conditional probability tables can be efficiently populated using maximum likelihood. Model structure learning was conducted using heuristic search procedures implemented in the SMILE modeling environment developed by the Decision Systems Laboratory at the University of Pittsburgh [16].

The hyperparameters for each model are the number of evenly spaced bins used in discretizing the continuous variables. The set of hyperparameters for each model, variable ranges, and the set of candidate bin counts considered for each

TABLE I: Hyperparmeter sets by model

Model	Variables Requiring Bin Counts
$\mathcal{M}_m$	v, d_{cl}, ϕ, d_{front}
$\mathcal{M}_b$	v, d_{cl}, ϕ, d_{front}
$\mathcal{M}_c$	v, d_{cl}, ϕ, d_{front}, d_{rear}, v_{rear}, v_{front}
$\mathcal{M}_h$	v, d_{cl}, ϕ, d_{front}, d_{rear}, v_{rear}, v_{front}, ρ_{scene}

TABLE II: Variable candidate bin counts

Model	Candidate Bin Counts
$\mathcal{M}_m$	3, 5, 10, 50, 100, 250 and 500
$\mathcal{M}_b$	3, 5, 7, 10 and 15
$\mathcal{M}_c$	3, 5, 7, 10 and 15
$\mathcal{M}_h$	3, 5, 7, 10 and 15

variable are listed in Table I, Table II, and Table III.

V. Evaluation

In this section we compare the performance of the scene distribution models. A model for each proposed method was trained on each scene set extracted from the three highway regions in the Highway 101 dataset and the two highway regions in the Interstate 80 dataset. The average cross-validated log-likelihoods of each model for each region are shown in Fig. 4.

Further validation can be performed by extracting emergent metrics from scenes generated from each model. The metrics include the mean vehicle counts, the mean time to collision for vehicles in a scene, and the ratios of lane speed and lane density between the right-most and left-most lanes in a given scene. The distributions over each metric variable were extracted from 1000 samples from models trained on the 101b dataset and are given in Fig. 5. The corresponding optimal model structure and variable bin counts are given in Fig. 6.

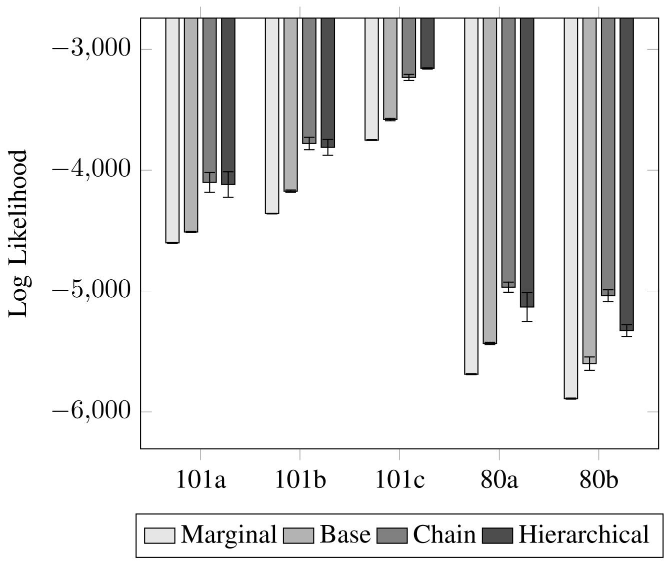


Fig. 4: Mean cross-validated log likelihoods of candidate models over selected sections, with 95% confidence intervals. Shorter bars indicate higher (better) likelihoods.

TABLE III: Variable ranges

Variable	Range		
v_{front}, v_{rear}, v	0	30.5	$\text{m}\,\text{s}^{-1}$
d_{front}, d_{rear}	0	91.5	m
d_{cl}	−1.75	1.75	m
ϕ	−0.1	0.1	rad
ρ_{scene}	0	60	vehicles

$$\text{extreme lane speed ratio} = v_{\text{rightmost}} / v_{\text{leftmost}} \tag{9}$$

$$\text{extreme lane density ratio} = \rho_{\text{rightmost}} / \rho_{\text{leftmost}} \tag{10}$$

The mean cross-validation likelihood score generally increases with model complexity. The chain model possesses the highest (best) mean cross-validation likelihood for four out of five road segments, with the hierarchical model outperforming it in section 101c. The mean cross-validation likelihood scores for the hierarchical model are similar to those for the chain model, but the hierarchical model typically possesses higher variance. This suggests that the addition of the global scene density and lane class variables may not be supported by the present data set sample size.

Distributions over the emergent metrics show rough consistency in terms of number of vehicles, with the marginal model producing the best mean and the chain model producing the best overall fit. The joint and hierarchical models produce the best matches for time to collision, but the true distribution has a slightly longer tail. All four models show similar behavior for the extreme ratios, indicating that the inclusion of the lane class in the hierarchical model presently does not have the desired effect.

VI. Conclusion

This paper introduces a methodology for learning distributions over highway scenes from real-world data for use in microscopic automotive simulation. Bayesian networks were used to represent distributions over vehicle positions, orientations, speeds, and other contextual variables for use in successively generating vehicles along a lane. The model parameters and structure are directly inferred from recorded data. Four models of increasing complexity were learned from the Next-Generation Simulation datasets for Highway 101 and Interstate 80 and their performance was assessed both qualitatively and quantitatively. The resulting models produce realistic driving scenes and can be sampled as many times as needed for automotive safety system validation and risk assessment through simulation.

Future work will include vehicle correlation across lanes. Models will support general road compositions such as merge ramps and urban features such as intersections, stoplights, and crosswalks. Finally, these models will be extended to distinguish between vehicle classes such as motorcyclists and buses, as these classes tend to exhibit different driving behaviors. Additional information and implementation details are available for download at https://github.com/sisl/HighwaySceneModel.

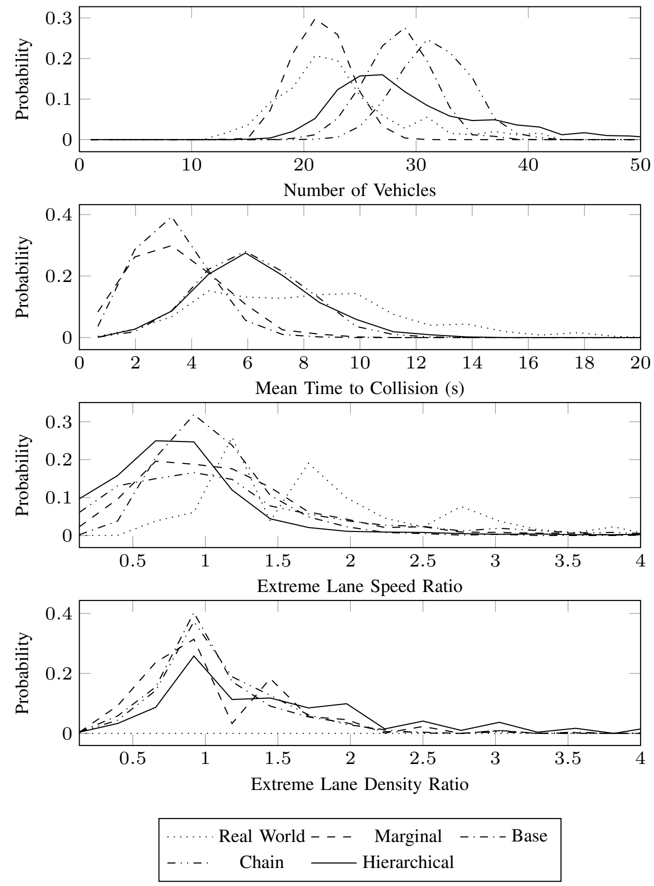


Fig. 5: Distributions over metrics from sampled scenes from Highway 101 section 101b. The extreme lane density ratio for real-world data approaches infinity due to the lack of vehicles in the auxiliary lane.

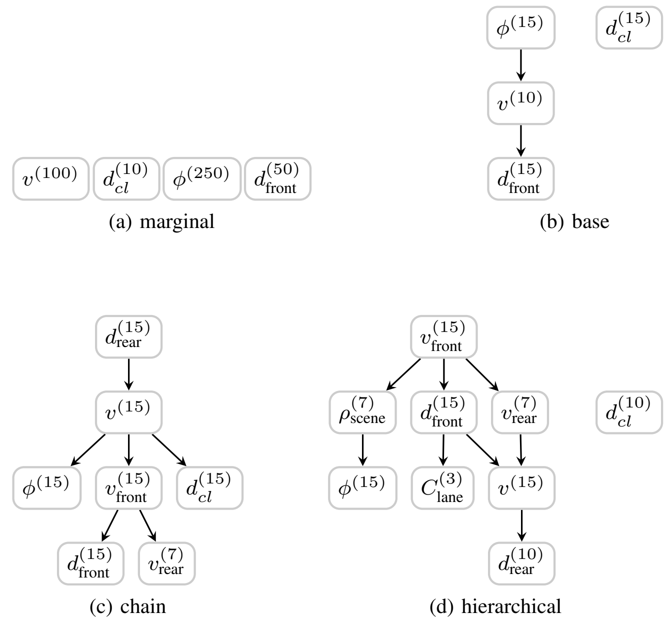


Fig. 6: Optimal model network structures for Highway 101 section 101b. Bin counts are listed in parentheses.

Acknowledgment

This work is sponsored by Robert Bosch LLC. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by Bosch. The authors appreciate the support and assistance provided by the Automated Driving team at Bosch. The authors gratefully acknowledge the helpful comments received from anonymous reviewers.

References

[1] M. J. Kochenderfer, M. W. Edwards, L. P. Espindle, J. K. Kuchar, and D. J. Griffith, "Airspace encounter models for estimating collision risk," *AIAA Journal on Guidance, Control, and Dynamics*, vol. 33, no. 2, pp. 487–499, 2010.

[2] M. J. Kochenderfer, J. E. Holland, and J. P. Chryssanthacopoulos, "Next-generation airborne collision avoidance system," *Lincoln Laboratory Journal*, vol. 19, no. 1, pp. 17–33, 2012.

[3] R. E. Chandler, R. Herman, and E. W. Montroll, "Traffic dynamics: studies in car following," *Operations Research*, vol. 6, no. 2, pp. 165–184, 1958.

[4] M. Brackstone and M. McDonald, "Car-following: a historical review," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 2, no. 4, pp. 181–196, 1999.

[5] X. Chen, L. Li, and Y. Zhang, "A Markov model for headway/spacing distribution of road traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 773–785, 2010.

[6] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "SUMO Simulation of Urban MObility," in *International Conference on Advances in System Simulation*, 2011.

[7] Q. Yang and H. N. Koutsopoulos, "A microscopic traffic simulator for evaluation of dynamic traffic management systems," *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 3, pp. 113–129, 1996.

[8] R. E. Neapolitan, *Learning Bayesian Networks*. Upper Saddle River, NJ: Prentice Hall, 2004.

[9] J. Colyar and J. Halkias, "US highway 101 dataset," Federal Highway Administration (FHWA), Tech. Rep. FHWA-HRT-07-030, Jan. 2007.

[10] ——, "US highway 80 dataset," Federal Highway Administration (FHWA), Tech. Rep. FHWA-HRT-06-137, Dec. 2006.

[11] C. Thiemann, M. Treiber, and A. Kesting, "Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data," *Transportation Research Records*, vol. 2088, pp. 90–101, 2008.

[12] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.

[13] D. M. Chickering, "Learning Bayesian networks is NP-complete," in *Learning from Data: Artificial Intelligence and Statistics V*, D. Fisher and H. Lenz, Eds., New York: Springer-Verlag, 1996, pp. 121–130.

[14] D. Heckerman, "A tutorial on learning with Bayesian networks," Microsoft Research, Redmond, WA, Tech. Rep. TR MSR-TR-95-06, 1995.

[15] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning Bayesian networks: the combination of knowledge and statistical data," *Journal of Machine Learning*, vol. 20, no. 3, pp. 197–243, Sep. 1995.

[16] M. J. Druzdzel, "SMILE: structural modeling, inference, and learning engine and GeNIe: a development environment for graphical decision-theoretic models," in *AAAI Conference on Artificial Intelligence (AAAI)*, Jul. 1999.