Modeling Multi-Vehicle Interaction Scenarios Using Gaussian Random Field

Yahouï Guo\textsuperscript{1}, Vinay Varma Kalidindi\textsuperscript{1}, Mansur Arief\textsuperscript{1}, Wenshuo Wang\textsuperscript{1}, Jiacheng Zhu\textsuperscript{1}, Huei Peng\textsuperscript{2}, Ding Zhao\textsuperscript{1,*}

Abstract— Autonomous vehicles (AV) are expected to navigate in complex traffic scenarios with multiple surrounding vehicles. The correlations between road users vary over the time, the degree of which, in theory, could be infinitely large, and thus posing a great challenge in modeling and predicting the driving environment. In this research, we propose a method to reproduce such high-dimensional scenarios in a finitely tractable form by defining a stochastic vector field model in multi-vehicle interactions. We then apply non-parametric Bayesian learning to extract the underlying motion patterns from a large quantity of naturalistic traffic data. We use Gaussian process to model multi-vehicle motion, and Dirichlet process to assign each observation to a specific scenario. We implement the proposed method on NGSim highway and intersection data sets, in which complex multi-vehicle interactions are prevalent. The results show that the proposed method is capable of capturing motion patterns from both settings, without imposing heroic prior, hence can be applied for a wide array of traffic situations. The proposed modeling can enable simulation platforms and other testing methods designed for AV evaluation, to easily model and generate traffic scenarios emulating large scale driving data.

I. INTRODUCTION

The deployment of an autonomous vehicle (AV) on public roads requires the AV be able to interact with naturalistic driving scenarios, most of which involve multiple road users. On the other hand, public expectation that an AV shall be able to drive and merge seamlessly in complex traffic keeps growing [1]. Hence, the modeling of multi-vehicle interaction scenarios is inevitable.

Traditionally, due to the limitations of the training data sets, researchers and engineers in transportation rely on strong assumptions to keep the inference tractable. Researchers often assume some fixed number of vehicles to be considered within a prediction window [2], consider all road users using the same driving strategies [3], or simplify the systems by only simulating one-to-one interactions [4], which restrict the applicability of such models for the study of naturalistic driving systems. To alleviate these assumptions, the model shall be able to consider the interactions among vehicles and avoid the restriction of presupposing the number of vehicles involved, which is a challenging task.

To fulfill the modeling requirements, we combine Gaussian Process (GP) with Dirichlet Process (DP). GP has been proven to be effective in modeling trajectory patterns. [5] uses Gaussian vector random field to model observed trajectories. [6] constructs a Gaussian random field model on fully or partially observed trajectories to perform classification and prediction. However, GP has rarely been used for modeling multi-agent interactions. In this study, we build a multi-vehicle interaction model by combining the effectiveness of GP [7] to model high-dimensional primitive patterns and the versatility of the DP [8] as a nonparametric method to model a stochastic system in a Bayesian view. By combining these methods, we are capable of modeling a highly dynamic multi-vehicle interaction, alleviating the need to pre-specify the number of vehicles involved, the driving events considered, and the heroic assumption about the independence between vehicles in the systems. Therefore, the motion patterns learned by the proposed model are fully data-driven, hence allowing us to characterize the observed traffic scenes and generate variants of the observed driving cases based on the learned motion patterns.

Many existing multi-vehicle modeling methods are focused on trajectory prediction. Some of these approaches employ neural network based methods while some others use stochastic models. [9] uses LSTM to model surrounding vehicles and predict their motion with the experiments limited to highway datasets. In [10], deep convolutional neural network is used to output predicted trajectories with associated probabilities. [11] uses context aware, Markovian models to describe multi-agent behavior and dynamic Bayesian networks to perform the prediction. [12] applies Gaussian process for multi-modal maneuver recognition and trajectory prediction using regression for likelihood calculations. In modeling motion patterns for trajectory prediction, [13] shows DP-GP performs better than Markov based models, but with the experiments limited to single vehicle. DP-GP modeling has also been used in the prediction of pedestrian trajectories [14].

While prediction can be seen as a popular outcome of modeling multi-agent motion using DP-GP, we study how it can also prove useful in capturing the interaction scenarios based on traffic data and generate simulated trajectories emulating the data.

This paper proceeds as follows. In Section II, the formulation of our proposed scheme is presented, followed by the overall framework in Section III. The experiment and results are described in Section IV and the findings are discussed in Section V. Finally, the conclusion and future work is summarized in Section VI.
II. FORMULATION

The multi-vehicle motion model is defined as an infinite Gaussian mixture [15] \( G \) of interaction scenarios with each mixture component defined by a Gaussian Process (GP) \( g_k \)

\[
G = \sum_{k=1}^{K} \pi_k g_k
\]

where \( \pi_k \) are the respective mixing proportions defined using Dirichlet Process (DP) prior. The problem can be therefore divided into several parts: how to define motion patterns \( g_k \), how to determine the number of mixtures \( K \), and how to infer the model parameters.

A. Gaussian process motion patterns

1) Modeling motion pattern with Gaussian process velocity vector field: We define a motion pattern \( g \) as a GP that maps from position coordinates to velocity, as shown in Fig. 1.

\[
g: (x, y) \rightarrow (v_x, v_y)
\]

where \( A_{ROI} \) is the region of interest.

![Gaussian process mapping from position to velocity vector field in A_{ROI}](image)

The velocity field information within a small region is expected to be consistent— a property that motivated the exploitation of GP to capture the consistency. A GP is defined as a collection of random variables, whose arbitrary subset has Gaussian distribution [7]. A GP motion pattern \( g \) here models the velocity \((v_x, v_y)\) as Gaussian random variables. Here, \( v_x \) and \( v_y \) are assumed to be independent for simplicity. For concise representation, we shall use indicator \( p \in \{x, y\} \) to avoid writing equations for both \( x \) and \( y \) directions— for example, the velocity \((v_x, v_y)\) at \((x, y)\) is written as \( v_p(x, y) \) or simply, \( v_p \).

A GP is fully specified by its mean function and covariance function:

\[
\begin{align*}
\mu_p(x, y) &= \mathbb{E}[v_p(x, y)] \\
\text{cov}_p(x, y) &= k_p((x, y), (x', y')) \\
&= \mathbb{E}[(v_p(x, y) - \mu_p(x, y))(v_p(x', y') - \mu_p(x', y'))]
\end{align*}
\]

and is written as

\[
v_p(x, y) \sim \mathcal{GP} (\mu_p(x, y), \text{cov}_p(x, y))
\]

where \((x', y')\) is any position coordinate in \( A_{ROI} \). The covariance \( k_p \) is defined using squared exponential function as follows:

\[
k_p(x, y, x', y') = \sigma^2_p \exp(-\frac{(x-x')^2}{2w_x^2} - \frac{(y-y')^2}{2w_y^2})
\]

where \( \sigma^2_p \) is the variance of the \( p \) component speed; \( w_x \) and \( w_y \) are the characteristic length-scale parameters, the inference of which is discussed in detail in section III.

The observed data is assumed to have an additive independent identically distributed Gaussian noise \( \varepsilon \) with zero mean and variance \( \sigma^2_{\varepsilon} \). As such, the covariance function for the noisy observed velocities \( v_p \) and \( v_p' \) becomes

\[
\text{cov}_p(x, y) = k_p((x, y), (x', y')) + \sigma^2_{\varepsilon}\delta((x, y), (x', y'))
\]

where \( \delta \) is the Kronecker delta function defined as

\[
\delta((x, y), (x', y')) = \begin{cases} 
1, & x = x' \text{ and } y = y' \\
0, & \text{otherwise}
\end{cases}
\]

Each observation \( s \), referred to as a frame, is a sample from a time-series data \( S \) and is a 2-dimensional representation of a given \( A_{ROI} \). Each frame is considered to have the position information \((x, y)\) and the corresponding velocity information \((v_x, v_y)\) of all the vehicles observed in \( A_{ROI} \) at that time. In vector form, the observed data is given by \( S = (V_x, V_y, X, Y) \). Here, \( [V_p]_j \) is the observed velocity at \(((x)_j, (y)_j)\) where \([x]_j\) denotes the \( j \)-th element of a vector. Here, \( [V_x]_j, [V_y]_j, [X]_j, [Y]_j \) are also vectors of the observed data of all agents in frame \( s \). Similarly, we write the testing data as \( S' = (V^*_x, V^*_y, X^*, Y^*) \), where \( V^*_p \) is unknown.

By definition of GP, the output has a joint Gaussian distribution given by:

\[
\begin{bmatrix}
V^*_p \\
V_p
\end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix}
\mu_p(X^*, Y^*) \\
\mu_p(X, Y)
\end{bmatrix}, \begin{bmatrix}
K_p(X^*, Y^*, X^*, Y^*) & K_p(X^*, Y^*, X, Y) \\
K_p(X, Y, X^*, Y^*) & K_p(X, Y, X, Y) + \sigma^2_{\varepsilon}I
\end{bmatrix} \right)
\]

where \( \mu_p(X^*, Y^*) \) and \( \mu_p(X, Y) \) are the prior distribution of the velocity. Conditioned on the observation \( V_p \), the posterior distribution of \( V^*_p \) is still Gaussian

\[
V^*_p | V_p, (X^*, Y^*), (X, Y) \sim N \left( \mu^*_p(X^*, Y^*), \text{cov}^*_p(X^*, Y^*) \right)
\]

where

\[
\begin{align*}
\mu^*_p(X^*, Y^*) &= \mu_p(X^*, Y^*) + K_p(X^*, Y^*, X, Y) \star \left[ K_p(X, Y, X, Y) + \sigma^2_{\varepsilon}I \right]^{-1} (V_p - \mu_p(X, Y)) \\
\text{cov}^*_p(X^*, Y^*) &= K_p(X^*, Y^*, X^*, Y^*) - K_p(X^*, Y^*, X, Y) \star \left[ K_p(X, Y, X, Y) + \sigma^2_{\varepsilon}I \right]^{-1} K_p(X, Y, X^*, Y^*)
\end{align*}
\]

2) Multi-vehicle trajectory generation from motion patterns: In order to calculate the likelihood of a GP motion pattern \( g \) given frame \( s = \{(x_j, y_j, v_{xj}, v_{yj}) \mid j = 1, 2, \ldots, l\} \), we need to specify how \( g \) generates \( s \). We model this procedure in three-steps: drawing the number of agents \( l \), the location of the agents \( \{(x_j, y_j) \mid j = 1, 2, \ldots, l\} \) and the velocity \( \{(v_{xj}, v_{yj}) \mid j = 1, 2, \ldots, l\} \) respectively.

Let \( a_m = \# \{y_i \mid l_i = m, y_i \in Y \} \) be the number of observations with \( m \) agents observed, \( m = 1, 2, \ldots \), and assign the weights \( a_m = a_m^0 / \sum_{q=1}^{a_{max}} a_q^0 \) to \( \delta(m) \), where \( \delta(m) \) is the distribution concentrated at a single point \( m \), the
empirical distribution \( \phi(Y) \) of the number of vehicles \( l \) is as follows:
\[
l \sim \phi(Y) = \sum_{q=1}^{\infty} a_{q}\delta(q)
\]

To draw the location of each car independently for the empirical population of cars over \( A_{ROI} \), we shall use a mixture of uniform distributions to fit the location distribution from \( S \). First, we discretize \( A_{ROI} \) into disjoint bins \( \mathcal{A} = \{ A_1, A_2, \ldots, A_n \} \), such that \( A_i \cap A_j = \emptyset \) for any \( i, j \in \{1, 2, \ldots, n\} \) and \( \cup_i A_i = A_{ROI} \); then we account the number of agents appeared in each bin \( a'_{Am} = \# \{ (x'_i, y'_j) \mid (x'_i, y'_j) \in A_m \cap y_i \in Y \} \), and assign a weight \( a_{Am} = a'_{Am}/\Sigma_{q=1}^{n} a'_{Aq} \) to each bin \( A_m \). We have:
\[
(x_j, y_j) \sim \psi(Y, A) = \sum_{q=1}^{n_A} a_{AQ} u(A_q), \ j = 1, 2, \ldots, l
\]
where \( u(A_q) \) is a uniform distribution over bin \( A_q \).

For sampling the velocity for each agent from \( g \), similar to the notation in (2) in which we use \( s = (V_x, V_y, X, Y) \), we have
\[
V_p \sim \mathcal{N}(\mu_p (X, Y), K_p (X, Y, X, Y)).
\]
The likelihood of motion pattern \( g \) given observation \( s \) therefore is
\[
p(s|g) = \phi(l; Y) \cdot \prod_{j=1}^{l} \psi((x_j, y_j); Y, A) \cdot \prod_{p \in \{ x, y \}} \mathcal{N}(V_p; \mu_p (X, Y), K_p (X, Y, X, Y))
\]
(6)

For a given dataset, the empirical distributions \( \phi \) and \( \psi \) are implicitly defined by the data. However, the discussed formulation enables the model to scale to data generation platforms where the distributions are expected to be explicitly defined.

B. Dirichlet Process Mixture of Motion Model

The proposed model considers the dataset \( \mathcal{S} = \{ s_i \mid i = 1, 2, \ldots, N \} \) as generated by an infinite mixture of motion patterns as shown in (4). Since the total number of the motion patterns \( K \) is not known, we give \( G \) a Dirichlet Process (DP) prior mixture weight. A DP is a distribution over distributions with infinite components. In our case, however, since the number of observations \( N \) is finite, only finitely many components will be discovered from the data.

An indicator variable \( z_i \) is introduced where \( z_i = k \) means the frame \( s_i \) is associated with latent motion pattern \( g_k \). The predictive distribution of \( g_{z_i} \), conditioned on the other motion patterns \( g_{z_{-i}} = \{ g_k \mid z_k \in z_{-i} \} \), where \( z_{-i} = \{ z_k \mid k = 1, 2, \ldots, n, k \neq i \} \), is
\[
g_{z_i} \mid g_{z_{-i}}, z_{-i} \sim \frac{1}{\alpha + N - 1} \left( \alpha G + \sum_{z_k \in z_{-i}} \Delta(g_{z_k}) \right)
\]
where \( \alpha \) is the concentration parameter and \( \Delta(g_{z_k}) \) is the point mass at \( g_{z_k} \). Then the prior probability of \( s_i \) belonging to an existing motion pattern \( g_k \) or an unseen motion pattern \( g_{K+1} \) is given by
\[
p(z_i = k \mid z_{-i}, \alpha) = \frac{n_k^{i-1}}{N - 1 + \alpha}, k = 1, 2, \ldots, K
\]
\[
p(z_i = K + 1 \mid z_{-i}, \alpha) = \frac{\alpha}{N - 1 + \alpha}
\]
(7)

where \( n_k \) is the number of observations currently assigned to \( g_k \) and \( n_k^{i} = \sum_{z_k \in z_{-i}} I[z_k = k] \).

Combining the likelihood from (6) and prior from (7), we have the posterior distribution of \( z_i \) as
\[
p(z_i = k \mid z_{-i}, \alpha, s_i, g_k) = \frac{n_k^{i-1}}{N - 1 + \alpha} p(s_i|g_k), \quad k = 1, 2, \ldots, K
\]
\[
p(z_i = K + 1 \mid z_{-i}, \alpha, s_i) = \frac{\alpha}{N - 1 + \alpha} \int p(s_i|g) \, dg
\]
(8)
The integration \( \int p(s_i|g) \, dg \) calculates the likelihood of all the motion patterns \( g \) over the base measure \( G \) given observation \( s_i \).

III. FRAMEWORK

In order to find a posterior motion pattern mixture, we use Gibbs sampling to infer the parameters of the model. For every iteration of Gibbs sampling, the model parameters and the mixture assignment of frames into motion patterns are updated.

A. Mixture model assignment

The assignment of all frames from \( \mathcal{S} \) is performed according to (8).

The likelihood \( p(s_i|g_k) \) for assigning frame \( s_i \) into existing pattern \( g_k \) defined in (6) is computed using the GP posterior from (4), with the training data now given by \( S^{g_k} = (V_{x_{g_k}}, V_{y_{g_k}}, X^{g_k}, Y^{g_k}) \), which is the vector form of the data of the \( n_k^{i} \) frames clustered under \( g_k \) and the testing data \( (V_{x_{s_i}}, V_{y_{s_i}}, X_{s_i}, Y_{s_i}) \) of frame \( s_i \)
\[
V_p \mid V_{g_k}, (X^{g_k}, Y^{g_k}), (X_{s_i}, Y_{s_i}) \sim \mathcal{N}(\mu'_p (X^{g_k}, Y^{g_k}), \text{cov}'_p (X^{g_k}, Y^{g_k}))
\]
Fig. 3. DP-GP mixture model simulation of traffic scenarios

where \(\mu^{\ast}_p\) and \(\text{cov}^\ast_p\) hold the same definition as in (5). A maximum a posteriori estimation is then performed across all the motion patterns \(g_k\) to identify the assignment \(z_i\)

For assignment of frame \(s_i\) under a new, unseen pattern \(g_{K+1}\), Monte-Carlo (MC) integration is used to approximate the likelihood integral \(\int p(s_i|g) \, dg\). Each MC iteration samples a new motion pattern using priors of model parameters and computes the likelihood using the GP prior given as

\[
V^s_p((X_s^*, Y_s^*) \sim GP(\mu_{p0}(X_s^*, Y_s^*), K_{p0}(X_s^*, X_s^*, Y_s^*, Y_s^*))
\]

where \(\mu_{p0}\) and \(\sigma^2_p\) are set to the data mean and variance respectively, and \(w_p\) is sampled using the prior defined later in (9).

B. Model parameters

The length scale parameters \(w_x\) and \(w_y\) from the exponential covariance calculation in (3) are given vague gamma prior

\[
w_p \sim \Gamma(a, b)
\]

where shape factor \(a\) and scale factor \(b\) are constants. The posterior calculation of \(w_p\) uses likelihood given by the GP prior of the frame data \(S^g_k\) assigned under motion pattern \(g_k\). The parameters \(w_p\) are therefore updated by re-sampling from the posterior given by

\[
w_p | g_k \sim \Gamma(a, b) \cdot p \left( V^S_p | (X_{S^g}, Y_{S^g}) \right)
\]

For the concentration parameter \(\alpha\), similar to [15], an inverse gamma prior is chosen and is updated by re-sampling from the posterior distribution given by

\[
p(\alpha | K, N) \propto \frac{\alpha^{K+\alpha}}{\Gamma(K+\alpha)} \exp\left(\frac{-1}{2\alpha}\right)
\]

The inference algorithm is summarized in Algorithm 1.

Algorithm 1 Inference

Initialization

\[
K \leftarrow 1
\]

\[
w_p \sim \Gamma(a, b)
\]

\[
\alpha^{-1} \sim \Gamma(1, 1)
\]

for Gibbs sampling iterations do

Update Mixture Assignment (1):

for frames \(i = 1, 2, \ldots, N\) do

for motion patterns \(k = 1, 2, \ldots, K\) do

\[
p(z_i = k | z_{i-1}, \alpha, s_i, g_k) = \frac{n^i_k}{n_{\alpha+1}^i} p(s_i | g_k)
\]

\[
p(z_i = K+1 | z_{i-1}, \alpha, y) = \frac{\int p(s_i | g) \, (dg)}{n^i_{\alpha+1}}
\]

end

Update \(z_i = \arg\max_k p(z_i = k | z_{i-1}, \alpha, s_i, g_k)\)

where \(k^+ = 1, 2, \ldots, K + 1\)

Update \(K, n_k\)

end

Update Model Parameters (1):

for motion patterns \(k = 1, 2, \ldots, K\) do

\[
w_p \sim w_p | g_k
\]

\[
\alpha \sim \alpha | K, N
\]

end

end

After the Gibbs sampling iterations, the posterior mixture model is used to generate simulated trajectories as shown in Fig 3. Given a test frame, a motion pattern assignment is
performed using this mixture which is used to generate the GP posterior mean velocity field. The velocity field defines the multi-vehicle trajectory simulation on the test frame.

IV. EXPERIMENT AND RESULTS

A. Experiment Setup

For evaluating the proposed motion model, a real world traffic dataset collected as part of Federal Highway Administration’s (FHWA) Next Generation SIMulation (NGSIM) project [16], [17], providing detailed multi-vehicle trajectory data as a time-series sequence, is chosen. The velocity information as \( v_x \) and \( v_y \) components is derived from this trajectory data. The model is evaluated on two traffic settings—highway dataset collected on a segment of the US Highway 101 (Hollywood Freeway) in Los Angeles, and intersection dataset collected on Lankershim Boulevard at Universal Hollywood Dr. in Los Angeles.

The inference algorithm is run for 100 Gibbs sampling iterations and is executed using parallel computing on a 44-core computer processor. The parameters \( a = 10 \) and \( b = 1 \) are chosen for gamma prior of the length scale parameters. The variance \( \sigma_n^2 \) for the additive Gaussian noise \( \epsilon \) is set to 0.04.

B. Highway Traffic Scenarios

The highway dataset is down-sampled to 1000 frames of time-sequence data with discretization of 0.5s. The mixture model resulted in 99 motion patterns being extracted from the data. The mixture proportion in a decreasing order is presented in Fig 4(a).

To generate the simulated multi-vehicle trajectories from the motion pattern results at the end of the Gibbs sampling iterations, a test frame is randomly chosen from the dataset. A motion pattern from the mixture is then assigned to the frame according to the assignment procedure discussed in section III. The derived mean GP velocity field is imposed on the vehicle distribution present in the \( A_{ROI} \) to run the simulation for that interaction scenario, and is presented in Fig 5. To illustrate the clustering, the original observations assigned under the same motion pattern are also included (with the vehicle velocity vectors shown) in the figure.

C. Intersection Traffic Scenarios

The results seen in the case of highway dataset are reproduced for the intersection dataset. The dataset is down-sampled to 600 frames of time-sequence data with discretization of 0.5s. The posterior mixture model consists of 86 motion patterns whose mixture proportion in a decreasing order is presented in Fig 4(b).

A simulated trajectory with the motion pattern vector field and the data observations are presented in Fig 6 similar to the results seen from the highway dataset.

V. DISCUSSION

A. Result analysis

The results from both the datasets demonstrate that the proposed model when applied to large time-series data sequences, extracts the underlying motion patterns which can be used to represent the interaction scenarios.

It can be noted in case of both the datasets that the frame indexes of the observation (training) frames clustered under the presented motion patterns are far apart in the data sequence. In the highway dataset, the frame indexes in 5(c) and 5(d) show that these observations are over 800 frames i.e. over 400s apart. The frame indexes from the intersection results in Fig 6 also show the same effect.

The motion pattern from highway dataset presented in Fig 5 indicates an interaction scenario with vehicles in some parts of the \( A_{ROI} \) (especially the rightmost lanes) moving faster than the others. This can be noted from the observations 5(c) and 5(d) and the same is also reflected in the simulated trajectory 5(b).

For the intersection based \( A_{ROI} \), the GP mean velocity field of the generated motion pattern presented in Fig 6(b) and the corresponding simulated trajectory 6(c) indicate a motion scenario where some of the vehicles oncoming from one direction are seen to take a left turn at the intersection while others continue straight with vehicles from the other direction standing still. This is a scene that the model generated on the test frame after learning the motion pattern.
The GP velocity field and simulated trajectory results show that the model learns the road physical layout from the data by exhibiting almost non-existent probability of the posterior vector field outside the road boundary. The model is also seen to learn lane information like which lanes correspond to a left turn, without having any explicit information of the road layout.

While the highway dataset has been clustered into motion patterns as expected, it offers little insights into the semantic visualization of the results due to the vehicle interactions being limited to motion in only one direction. In that sense, the intersection results involving vehicles’ motion in multiple directions provide further motion patterns generated from the intersection dataset are presented in Fig 7. While patterns 7(a) and 7(d) present the interaction scenarios with vehicles from both sides travel straight, 7(b) almost exclusively captures the left turning motion of the vehicles from one direction. 7(d) presents a more complex scenario with vehicle motion flow in many directions.

In the mixture proportion presented in Fig 4(a), the tail end of the plot indicates few motion patterns generated from only one frame i.e., each of those frames clustered as a single motion pattern. From the intersection results in Fig 4(b), there exist fewer such patterns in proportion to the total number of patterns. This could be due to the larger speeds of the vehicles on the highway which strongly affects the correlation between sequential frames due to the updated vehicle distribution in the \( A_{ROI} \). In the case of the intersection, the relatively lower speeds result in having the sequential frames more likely to be clustered under the same motion pattern, possibly in addition to other frames elsewhere in the time-sequence.

B. Limitations

The primary limitation of this work lies in the inference of the mixture model using Gibbs sampling. A termination criteria is not explicitly available, especially because of the unsupervised nature of the problem, due to which it is difficult to come up with a suitable number of iterations. An evaluation of the resultant mixture assignment could be defined to determine the convergence.

Furthermore, in a dataset involving a more complex intersection, modeling all the multi-vehicle interactions with a single Gaussian process might overly marginalize the true velocity information of the data. Future research shall employ multiple Gaussian processes possibly conditional on vehicle direction of motion and other information to model each interaction scenario. Also, the true road layout information is not included in the model. Although the results show that the model has learnt these boundaries from the data, the generated simulation trajectories can be treated with higher confidence if the road boundary and other traffic rules based information are embedded into the model wherever available.

VI. CONCLUSION AND FUTURE WORK

In this work, we formulate a model for multi-vehicle interaction scenarios using Gaussian process, a mixture of which
is generated from naturalistic data by using non-parametric Bayesian learning. By employing Dirichlet Process as the mixture of the motion model, we are able to alleviate the restriction on the number of motion patterns existing in the dataset, allowing the model to be fully data-driven. The experiment results using NGSim datasets demonstrate the extracted multi-vehicle interactions as motion patterns, capable of capturing the highly dynamic scenes from highways and intersections. This result allows modelers to extract multi-vehicle interaction scenarios efficiently from large-scale data, which can further be used for simulating complex traffic scenes, predicting the trajectories of vehicles in multi-vehicle systems, and evaluating the safety of AV when interacting with human driven vehicles in a complex driving situations.

ACKNOWLEDGMENT

Toyota Research Institute (TRI) provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity.

REFERENCES

[1] J. Claybrook and S. Kildare, “Autonomous vehicles: No driver no regulation?” Science, vol. 361, no. 6397, pp. 36–37, 2018.
[2] N. Deo, A. Rangesh, and M. M. Trivedi, “How would surround vehicles move? a unified framework for maneuver classification and motion prediction,” IEEE Transactions on Intelligent Vehicles, vol. 3, no. 2, pp. 129–140, 2018.
[3] M. Treiber, A. Hennecke, and D. Helbing, “Congested traffic states in empirical observations and microscopic simulations,” Physical review E, vol. 62, no. 2, p. 1805, 2000.
[4] D. Zhao, H. Lam, H. Peng, S. Bao, D. J. LeBlanc, K. Nobukawa, and C. S. Pan, “Accelerated evaluation of automated vehicles safety in lane-change scenarios based on importance sampling techniques,” IEEE transactions on intelligent transportation systems, vol. 18, no. 3, pp. 595–607, 2017.
[5] M. Barão and J. S. Marques, “Gaussian random vector fields in trajectory modelling,” in Irish Machine Vision and Image Processing Conference-IMVIP, 2017.
[6] K. Kim, D. Lee, and I. Essa, “Gaussian process regression flow for analysis of motion trajectories,” in 2011 International Conference on Computer Vision. IEEE, 2011, pp. 1164–1171.
[7] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. The MIT Press, 2006.
[8] Y. W. Teh, “Dirichlet proc.” Encyclopedia of machine learning, vol. abs/1804.10467, pp. 280–287, 2010.
[9] N. Deo and M. M. Trivedi, “Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms,” 05 2018.
[10] H. Cui, V. Radosavljevic, F. Chou, T. Lin, T. Nguyen, T. Huang, J. Schneider, and N. Djuric, “Multimodal trajectory predictions for autonomous driving using deep convolutional networks,” CoRR, vol. abs/1809.10732, 2018.
[11] J. Schulz, C. Hübmann, J. Löchner, and D. Burschka, “Interaction-aware probabilistic behavior prediction in urban environments,” CoRR, vol. abs/1804.10467, 2018. [Online]. Available: http://arxiv.org/abs/1804.10467.
[12] Q. Tran and J. Firl, “Online maneuver recognition and multimodal trajectory prediction for intersection assistance using non-parametric regression,” in 2014 IEEE Intelligent Vehicles Symposium Proceedings, June 2014, pp. 918–923.
[13] J. Joseph, F. Doshi-Velez, A. S. Huang, and N. Roy, “A bayesian nonparametric approach to modeling motion patterns,” Autonomous Robots, vol. 31, no. 4, p. 383, 2011.
[14] Y. F. Chen, M. Liu, S.-Y. Liu, J. Miller, and J. How, “Predictive modeling of pedestrian motion patterns with bayesian nonparametrics,” 01 2016.
[15] C. E. Rasmussen, “The infinite gaussian mixture model,” Advances in Neural Information Processing Systems, vol. 12, pp. 554–560, 2000.