Bayesian Nonparametric Modeling of Driver Behavior using HDP Split-Merge Sampling Algorithm

Vadim Smolyakov\textsuperscript{1}, and Julian Straub\textsuperscript{2} and Sue Zheng\textsuperscript{3} and John W. Fisher III\textsuperscript{4}

Abstract—Modern vehicles are equipped with increasingly complex sensors. These sensors generate large volumes of data that provide opportunities for modeling and analysis. Here, we are interested in exploiting this data to learn aspects of behaviors and the road network associated with individual drivers. Our dataset is collected on a standard vehicle used to commute to work and for personal trips. A Hidden Markov Model (HMM) trained on the GPS position and orientation data is utilized to compress the large amount of position information into a small amount of road segment states. Each state has a set of observations, i.e., car signals, associated with it that are quantized and modeled as draws from a Hierarchical Dirichlet Process (HDP). The inference for the topic distributions is carried out using HDP split-merge sampling algorithm. The topic distributions over joint quantized car signals characterize the driving situation in the respective road state. In a novel manner, we demonstrate how the sparsity of the personal road network of a driver in conjunction with a hierarchical topic model allows data driven predictions about destinations as well as likely road conditions.

I. INTRODUCTION

Vehicles are equipped with an increasing number of sensors and electronics to react dynamically to changing road conditions and to increase driver safety. As a result, large volumes of driver-specific data related to driving conditions and driver behavior are generated. We are interested in analyzing this data to learn models of driving behavior. Such models could be used to anticipate dangerous situations, to improve the driving schedule of a person, and to tailor various aspects of the driving experience to the individual.

Here, we use data collected from one vehicle’s sensors over numerous trips to construct a Hierarchical Dirichlet Process (HDP) model of driving behavior and road conditions. HDPs are commonly used for topic modeling of text corpora \cite{1}, \cite{2}, \cite{3}, \cite{4} to uncover the set of topics that comprise each document in the corpus. In our case, the documents are road segments and the words are associated quantized sensor measurements. The topics in the HDP model are sensor distributions in the road segments; these distributions capture the driving conditions in each road segment as encountered by the driver as well as their driving behavior and common driving conditions. To our knowledge this is a new approach for modeling driving behavior. Unlike related work which is based on assumptions about the capabilities and behaviors of humans (i.e., see for an overview \cite{5}), our model is purely data driven.

It is important to note that the hierarchy within the HDP model allows sharing of measurements across similar road segments. This is an appealing aspect of the model since it enables us to learn an expressive model for road segments which are visited rarely via similar road segments that are visited more often. In order to utilize an HDP model, we first organize the sensor data into "documents" (i.e., road segments and their associated quantized measurements). We consider the case in which a road map is not available, however, it is straightforward to incorporate such information. Additionally, typical drivers often traverse a small subset of the roads in the road network. We use a Hidden

\textsuperscript{1}Vadim Smolyakov is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA vss@csail.mit.edu

\textsuperscript{2}Julian Straub is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA jstraub@csail.mit.edu

\textsuperscript{3}Sue Zheng is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA zhengs@mit.edu

\textsuperscript{4}John W. Fisher III is with the Faculty of the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA fisher@csail.mit.edu

This work was partially funded by the Ford-MIT alliance and by the Office of Naval Research Multidisciplinary Research Initiative (MURI) program, award N00014-11-1-0688.
Markov Model (HMM) to learn the road segments. The HMM condenses position information from recorded trips into road segment states. The set of hidden states effectively corresponds to a sparse road network which consists only of the roads which the driver has traversed. We then use the trained HMM to associate sensor measurements to road segments to produce "documents" for the HDP model.

In addition to organizing the data for the HDP model, the HMM also provides insight into driver behavior such as typical routes and probable destinations. Special hidden states are introduced in the HMM to represent starting locations (sources) and destinations. Consequently, identification of the most likely route between two states and finding the distribution over probable destinations become well-posed questions and allow us to make route and destination predictions.

The contributions of this paper are (1) to show how sparsity in the HMM transition matrix together with starting and absorption states lead to accurate long term predictions of driver routes and destinations and (2) the novel application of a HDP split-merge sampler to model the joint distribution of quantized vehicle signals scalable to a large number of road-segments.

II. HIDDEN MARKOV MODEL

An HMM is used to model the trips that a driver takes through a road network. We explore two models for the HMM. In the first model, the hidden state corresponds to a road segment, a start location, or a destination. In this model, the future path is independent from the past path when conditioned on the current road segment. We expect this to be a poor model of driver behavior since this is likely an oversimplification; the past can provide considerable information about the future. For instance, drivers often do not return to a previously visited state within a trip (unless they are lost).

In our second model, we attempt to capture more of the trip history in the current state by augmenting the road states with the start location. Under this model, the road segment at the next time instance depends only on the current road segment and the start location. We will show that this model is more representative of driver behavior and provides accurate predictions of destinations and routes. We describe this second model below. The first model is a simplification of the described model.

A. Hidden States

Each hidden variable, \( x_t \), in the HMM (see Fig. 3) takes on a value from the set of hidden states, \( \mathcal{X} \). Source states, \( \mathcal{X}_S \), destination states, \( \mathcal{X}_D \), and road segment states augmented by the source state, \( \mathcal{X}_R \times \mathcal{X}_S \), compose the set of hidden states: \( \mathcal{X} = \mathcal{X}_S \cup \mathcal{X}_D \cup (\mathcal{X}_R \times \mathcal{X}_S) \). Destination states are absorbing states which are indicated by key-off events in the data. Similarly, source states are indicated by key-on events. The distribution over the initial state, \( x_0 \in \mathcal{X} \), is parameterized as \( p(x_0 = m) = \theta_m \). Conditioned on the current state, \( x_t \), the distribution for the next state, \( x_{t+1} \), is parameterized as

\[
p(x_{t+1} = m|x_t = k) = \theta_{km}.
\]

Since physically realizable transitions occur only between road segments in close proximity, we would expect most transition probabilities to be zero. We use a Dirichlet prior on the parameters with \( \alpha < 1 \) to favor a sparse transition matrix:

\[
p(\theta_{k1}, \theta_{k2}, \ldots, \theta_{k|\mathcal{X}|}) \propto \prod_{i=1}^{|\mathcal{X}|} \theta_{ki}^{\alpha-1}.
\]

B. Observation Model

Each trip contains measurements of position, \( r_t \) and heading, \( h_t \). When the vehicle has GPS, the recorded position is the GPS position; otherwise the reported position is obtained by dead reckoning, a process which estimates position by combining the previous position with aggregated incremental changes in a relative coordinate system. Positions which are inferred using dead reckoning are indicated by an inferred-position indicator, \( q_t \); for such measurements, we model a larger uncertainty associated with the measurement.

In addition to position and heading measurements, there is a key-on event at the start of each trip which indicates that hidden state must be from the set of source states. In the measurement model, we have a binary key-on indicator variable, \( k_t^{on} \), which takes value 1 if a key-on event occurs at time \( t \). Similarly, there is a key-off event and corresponding indicator variable, \( k_t^{off} \), which indicates that the hidden state is from the set of destination states. This set of measurements

---

**Fig. 2:** (Left) A standard HMM where the hidden variables (in red) correspond to road segments and the observation variables (in blue) include position and heading measurements. (Right) A conceptual rendering of the HMM. The physical road is shown in gray while the HMM representation of the road states is shown in red. Position measurements are shown in blue.
conditioned on the hidden state, the measurement model is as follows:

\[ p(r_t, h_t, q_t, k^{off}_t, k^{on}_t | x_t) = p(r_t | q_t, x_t) p(h_t | x_t) p(q_t | x_t) \times p(k^{off}_t | x_t) p(k^{on}_t | x_t). \]

Note that the conditional distribution for position depends on the value on the inferred-position indicator; a larger uncertainty is associated with the position when the position has been inferred. Position is Gaussian with state-dependent parameters:

\[ p(r_t | q_t, x_t) = \begin{cases} N(r_t; \mu_{r,x_t}, \Sigma_{r,x_t}) & q_t = 0 \\ N(r_t; \mu_{r,x_t}, c \cdot \Sigma_{r,x_t}) & q_t = 1 \end{cases} \]

where \( c > 1 \) is a constant used to capture the increase in uncertainty of the inferred position.

Heading is also Gaussian with its own state-dependent parameters:

\[ p(h_t | x_t) = N(h_t; \mu_{h,x_t}, \Sigma_{h,x_t}). \]

The inferred position indicator has a Bernoulli distribution with parameter, \( p_{z_t} \). The key-on and key-off measurements are indicators of source and destination states respectively: \( k^{on}_t = 1(x_t \in X_S) \), \( k^{off}_t = 1(x_t \in X_D) \) and have degenerate distributions.

We would expect measurements arising from the same physical location (road segment) to have parameters which do not depend on the source state. Therefore, the measurement parameters are independent of the source state when conditioned on the road segment state. That is, \( \mu_{r,x_t} = \mu_{r,x} \) for \( x_t \in \{x_r \times X_S\} \), where \( x_r \in X_R \), and likewise for the other measurement parameters. When we estimate the measurement parameters for a road segment, this formulation allows us to aggregate observations from trips which start at different locations but share this physical road.

Similarly, there will be pairs of source and destination states which correspond to the same physical location. If a trip ends at a given location, the next trip will typically start from the same location. Since this pair of states share physical properties, these states will share measurement parameters.

C. EM Updates

Given the volume of data under consideration, we find that an EM formulation using explicit state assignments provides a tractable learning approach. This approach yields locally optimal values for the set of parameters \( \psi = \{ \theta_m, \theta_{km}, \mu_{r,m}, \Sigma_{r,m}, \mu_{h,m}, \Sigma_{h,m}, p_m \} \) for \( m, k \in \mathcal{X} \) using measurements from \( N \) trips. The EM updates consist of iteratively finding the most likely assignment for the hidden states given previous parameter estimates, then using these assignments to improve the parameter estimates. The reader is referred to [6] for an introduction to the EM algorithm.

To initialize the parameters, we run DP means [7] to cluster the measurements based on position and heading; a state is created from each cluster. DP means allows us to initialize the model without pre-specifying the number of states. Measurements assigned to the cluster (state) are used to calculate initial values for measurement model parameters. The transition matrix is initialized as a full matrix with higher probability for states which are closer together. The distribution for the first state is initialized as a uniform distribution.

D. Predicting Routes and Destinations

Using the HMM model, we can predict a driver route from state \( a \) to state \( b \) by identifying the sequence of states \( \{x_1^t, \ldots, x_{T-1}^t, x_T^t\} \) that maximizes the probability:

\[ p(x_1^t, \ldots, x_{T-1}^t, x_T^t | y_1, \ldots, y_T, \theta) \]

subject to the transition and measurement models. This can be done using the Viterbi algorithm, which efficiently computes the most likely sequence of states given the observations.

Fig. 3: Predicted routes and absorption probabilities corresponding to the five most likely destinations of the model without (left) and with (right) start locations. The most likely destinations differ between the two models; notably, in the left model, there is significant probability that the driver will return to his starting location (shown in yellow) at the start of the trip, while the starting location is not a likely destination in the model augmented with start location. Furthermore, the absorption probability of the true destination, shown in green, dominates over alternative possible destinations sooner in the bottom model.
the reader is referred to the original papers [1] and [2].

cretized car measurements. For a more detailed presentation
sampler to perform inference on a large number of dis-
and describe how we use a parallelizable HDP split-merge

F. HDP model

This means that words are described by a multidimensional
measure equivalent to the classical text corpus topic model.

Augmenting the sample space with sub-clusters leads to
proposals of likely splits and merges. A combination of a
parallelizable split-merge HDP sampler described in [2].

G. Split-Merge HDP Sampler

Because of the large quantity of data, we use highly
parallelizable split-merge HDP sampler described in [2].
Augmenting the sample space with sub-clusters leads to
proposals of likely splits and merges. A combination of a
restricted Gibbs sampler (that does not create new clusers)
with split/merge moves results in an ergodic Markov chain.

1) Restricted Gibbs Sampler: Let $m_{jk}$ denote the number
of clusters in document $j$ with shared topic $k$ and let $n_{jtk}$
denote the number of words in document $j$ in cluster $t$ with
topic $k$. Then, the marginal counts $n_{jt} = \sum_{t,k} n_{jtk}$ and
$m_{j} = \sum_{k} m_{jk}$ represent the number of words and topics
in document $j$, respectively. Extending the DA sampling
algorithm results in the following restricted posterior dis-

\[
\prod_{i=1}^{N-1} p(x_{i+1}^a|x_i^a) = \max_{n} \max_{x_1,\ldots,x_n} \prod_{i=1}^{n-1} p(x_{i+1}|x_i) \prod_{i=1}^{N-1} p(x_i^a)
\]

It is well known, [8], that this can be formulated as a shortest
path problem by defining a graph on the hidden states with
edge weights $w_{ij} = -\log p(x_j|x_i)$. Additionally, from any road segment state, $i$, we can
find the probability of reaching any destination state, $j$;
this is known as the absorption probability. The absorption
probability, $a_{ij}$, is the probability of reaching absorbing state,
$j$, if the chain starts from state, $i$, and can be found by solving
the following set of equations:

\[
a_{jj} = 1 \quad \forall j \in \mathcal{X}_D \\
a_{ji} = 0 \quad \forall j \in \mathcal{X}_D \quad \forall i \neq j \\
a_{ij} = \theta_{ij} + \sum_{k \in \mathcal{X} \setminus j} \theta_{ik} a_{k_j} \quad \forall j \in \mathcal{X}_D, \quad \forall i \in \mathcal{X}_R
\]

This gives us a probability distribution over destinations
when we start from a given road state.

E. Bayesian Nonparametric Topic Modeling of Car Signals

Thus far, we have formulated an HMM model for driving
behavior which has predictive aspects. The model is also
used to organize the data into "documents" so that we can
perform HDP topic modeling on the dataset. In this section
we discuss how we combine a standard HDP model with the
use of an HMM to discover documents. We also relate our
HDP model for car signals to the classical HDP topic model.

To bridge the gap between the classical HDP topic modeling
of text corpora and the modeling of car signals, such as
velocity, acceleration and rotational speed, note the following
correspondences:

word $\leftrightarrow$ car signals at one instance in time

document $\leftrightarrow$ road segment

corpus $\leftrightarrow$ map

Each learned road segment from the HMM is used as
a document in the HDP model. To obtain a set of sensor
measurements associated with a road segment, or a set of
words from the document, we perform ML assignment of
road states for the trips and assign the corresponding sensor
measurements to those road states.

Since the car signals are continuous quantities, we quan-
tize them using DP means [7] and use a discrete base
measure equivalent to the classical text corpus topic model.
This means that words are described by a multidimensional
vector, which amounts to modeling the joint distribution over
all signals.

In the next section we briefly summarize the HDP model
and describe how we use a parallelizable HDP split-merge
sampler to perform inference on a large number of dis-
cretized car measurements. For a more detailed presentation
the reader is referred to the original papers [1] and [2].
Fig. 5: Plots in the first row show the topic assignments (left) and global topic distribution (right). The plots in the second row show held-out word (HOW) log-likelihood (left) and number of topics vs iterations (right). Finally, the top 10 likely topics inferred by HDP split-merge sampler are shown at the bottom row for different car signals.
tructions:
\[
\begin{align*}
\pi_j | \beta, z & \propto \sum_{k=1}^{K} \pi_{jk} f_x(x_{ji}; \theta_k) [z_{ji} = k] \quad (4) \\
\pi_j | \beta, z & \propto \sum_{k=1}^{K} \pi_{jk} f_x(x_{ji}; \theta_k) [z_{ji} = k] \quad (4) \\
\end{align*}
\]

Since \( p(\beta|m) \) is not known analytically, we use the auxiliary variable \( m_{jk} \), \( s(n,m) \) denotes unsigned Stirling numbers of the first kind. Note that the last components \( \beta_{K+1} \) and \( \pi_{j(K+1)} \) aggregate the weight of all empty topics. Finally, \( I_k = \{ j; i; z_{ji} = k \} \) denotes the set of indices in topic \( k \), and \( f_x \) and \( f_\theta \) denote the observation and prior distributions. The equations above can be sampled in parallel and fully specify the restricted Gibbs sampler.

The method combines a Gibbs sampler that is restricted to non-empty clusters with a Metropolis-Hastings (MH) algorithm that proposes splits and merges.

2) Subcluster Splits and Merges: For each topic \( k \), we fit two sub-topics \( kl \) and \( kr \) referred to as the left and right sub-clusters. Each topic is augmented with global sub-topic proportions \( \beta_k = \{ \beta_{kl}, \beta_{kr} \} \), document-level sub-topic proportions \( \pi_{jk} = \{ \pi_{jkl}, \pi_{jkr} \} \), and sub-topic parameters \( \theta_k = \{ \theta_{kl}, \theta_{kr} \} \). Moreover, each word \( x_{ji} \) is associated with sub-topic assignment \( z_{ji} \in \{ l, r \} \). Then the marginal posterior distributions can be derived [2] as:
\[
\begin{align*}
p(\beta_k | \cdot) & = \text{Dir}(\gamma + m_{kl}, \gamma + m_{kr}) \quad (6) \\
p(\pi_{jk} | \cdot) & = \text{Dir}(\alpha \beta_{kl} + n_{jkl}, \alpha \beta_{kr} + n_{jkr}) \quad (7) \\
p(\theta_{kh} | \cdot) & \propto f_s(x_{khi}; \theta_{kh}) f_\theta(\theta_{kh}; \lambda) \quad (8) \\
p(\bar{z}_{ji} | \cdot) & \propto \pi_{jkl} f_x(x_{ji}; \theta_{kl}) [z_{ji} = k] \quad (9) \\
p(m_{jkh} | \cdot) & = f_m(m_{jkh}; \alpha \beta_{kh}, \bar{n}_{jkh}) \quad (10) \\
\end{align*}
\]

Notice the similarity between these equations and ones derived earlier. Inference is performed by interleaving the sampling equations (1) – (5) with marginal posterior equations (6) – (10).

3) Metropolis-Hastings: A Metropolis-Hastings framework proposes splits and mergers of sub-clusters and either accepts or rejects them. Let \( v = \{ \beta, \pi, z, \theta \} \) and \( \bar{v} = \{ \bar{\beta}, \bar{\pi}, \bar{z}, \bar{\theta} \} \) be a set of regular and auxiliary variables, respectively. Then a sampled proposal \( \{ v, \bar{v} \} \sim q(\bar{v}, v|v) \) is accepted with probability:
\[
P_a = \min \left[ 1, \frac{p(x,v)q(\bar{v}|x,v)}{p(x,v)q(\bar{v}|x,v)} \right] \quad (11)
\]

H. Performance Evaluation

To evaluate the performance of the HDP model, we are computing the average log predictive probability of held-out words. To compute this probability, we split a test document into two sets: held out words \( w^{ho} \) and observed words \( w^{obs} \). Then we update the model using the observed words. This gives us the posterior parameters \( \{ \phi^{obs}, \lambda^{obs} \} \) for the test document which we in turn use to find \( \phi^{ho} \) for the held-out words. Now we can compute the probability of a held-out word as given all training data \( D \) as well as the observed words \( w^{obs} \) in this document:

\[
p(w_i | D, w^{obs}) = \sum_i q(z_i = i | \phi^{ho}) \sum_{k=1}^K q(c_i = k | \phi^{obs}) p(w_i | z_i = i, c_i = k, \lambda^{obs}) \quad (12)
\]

where \( p(w_i | z_i = i, c_i = k, \lambda^{obs}) = \frac{\lambda_{n_i}^{obs}(w_i)}{\sum_{w} \lambda_{n_i}^{obs}(w)} \) is the conditional distribution of a held-out word under the posterior distribution of words in this document.

As a model to compare the HDP to, we utilize a non-hierarchical model that assumes a Categorical distribution with a Dirichlet prior for the words in each road-state. These distributions are modeled completely independent – not connected via a hierarchy like in the HDP model. This allows us to compute posterior Categorical distributions given the observed words in each road-state.

III. RESULTS

In the following we will first give results for the predictive power of the HMM model before we describe a topic model for the joint distribution of speed and time-of-day measurements.

A. Dataset Description

Our dataset comprises of 1K trips recorded from a standard car used by a single driver. The routes are mostly commuting to work but also some longer range trips outside the city.

The GPS position and heading measurements of the car are used to train the HMM model. From various other signals of the car we selected quantized car velocity and time of day for the HDP topic model. These were selected, since they contain interesting information both about the driving behavior as well as the driving situation in a road state.

B. Predicting Routes and Destinations

To evaluate the quality of the learned HMM, we examine the ability of the HMM to predict the destination for 20 held-out trips under the two different models. Additionally, we compare the path of the held-out trips against the most likely route obtained from the transition matrix for the HMM.

Fig. 3 shows the performance of the two models on a held-out trip. The plots under the maps in the figure show the absorption probabilities for the probable destinations as a function of time. The maps above show the trip, the most
likely route between the source and destination state, and the locations of the probable destinations. While the most likely path between source and destination from both models agrees with the observed trip trajectory, we observe that the augmented model is able to identify the correct destination sooner than the first model. In fact, the first model is able to correctly predict the destination after 10% of the trip for only 3 of the held-out trips while the augmented model is able to do so for 11 of the trips.

In Fig. 6, we show the most likely destination for each road segment. For the augmented model, since each state associated with a road segment also has a start location, we’ve chosen a particular start location to illustrate the differences between these two models. In particular, when starting from the specified start location, we see that trips which traverse beyond destination 7 in Fig. 6 (bottom) are more likely to terminate at a destination which is further from the starting location. The unaugmented model is unable to make this distinction, so trips which traverse road segments near destination 5 in Fig. 6 on the top (which corresponds to destination 7 in Fig. 6 on the bottom) are likely to terminate at that destination.

The results show that the most likely route obtained by the models frequently align exactly with the path of the held-out trips. This can be explained through the sparsity of the transition matrices; since each state can only transition to few states, and very often just one state, long term predictions in this model are quite accurate.

C. HDP Model

We are quantizing velocity and time-of-day measurements to words that can be fed into the HDP inference algorithm. Quantization is performed via DP k-means clustering over the individual signals.

The speed measurements arrive at a rate of 1 Hz from the GPS sensor. There are 696k joint observations – velocity/time-of-day pairs – across all 12k road states. As can be seen in Fig. 1, these observations are distributed non-uniformly – we get a lot of measurements on the daily commute route and few on highways leading outside the city. This means that the road-state corpus has very imbalanced document sizes when compared to text corpus modeling. However, our results demonstrate, that this presents no issue to the inference algorithm.

We empirically found the following set of parameters: $\gamma = 10.0$ and $\alpha = 0.1$, corresponding to the global and local concentration parameters, respectively.

Fig. 5 demonstrates that the hierarchy in the HDP is able to pool measurements from different road-states to obtain a descriptive topic for these. For each road state we obtain the maximum likelihood (ML) topic assignment and plot the respective road states in red. This pooling of observations can for example be observed for topics 0 and 41, which consist of almost all highway road states as can be seen in the ML topic assignment plots (compare the red road segments to the highways depicted in map in Fig. 1).

Using the inferred mixture of topics for each state, we can now compute the ML estimate of the marginals for the individual sensor signals and plot them color-coded for each road state. Fig. 7 shows this for the marginal over speed and time-of-day. Comparing the spatial distribution of the ML speed estimates computed from the inferred HDP model in Fig. 7d with ML estimates obtained from the empirical distribution depicted in Fig. 7b, we can see that the HDP model is able to capture the distribution of the input data. Additionally, it is clear from the spatial distribution of the ML estimates of driving speeds, that the HDP model captures for example the fact, that inner city driving is slower than highway driving. The ML estimates of time-of-day (Fig. 7c and 7a) show that the trips outside the city were not undertaken in the morning or evening.

IV. CONCLUSION

We have shown that the inherent sparsity of the learned personal road network allows accurate long term predictions of driver routes. Additionally, augmenting the model with start location yields a more representative model which provides better destination predictions. Exploiting the hierarchy of the HDP topic model, we are able to learn expressive
Fig. 7: Plots of the Maximum Likelihood (ML) estimates of time-of-day and speed computed from the inferred HDP model (bottom row) and the empirical distribution (top row).

topic distributions despite the fact that the number of car signal measurements differs widely between different road states. The combination of both types of of models allows us to model the driving behavior of an individual driver. This type of model can for example assist in optimizing the daily commute route or help predict traffic jams. As a next step it would be interesting to compare the driver models for different drivers to allow driver classification based on the driving behavior.

ACKNOWLEDGMENT

We wish to thank Tom Pilutti and Shane Elwart for their help with acquiring the data and for insightful discussions regarding the methodology.

REFERENCES


