

Driver Behavior Classification in Crash and Near-Crash Events Using 100-CAR Naturalistic Data Set

Abdalla Abdelrahman
School of Electrical and
Computer Engineering
Queen's University
Ontario, Kingston
Email: a.abdelrahman@queensu.ca

Najah Abu-Ali
College of Information Technology
United Arab Emirates University
Al-Ain, UAE
Telephone: +97137135543
E-mail: najah@uaeu.ac.ae

Hossam S. Hassanein
School of Computing
Queen's University
Ontario, Kingston
Email: hossam@cs.queensu.ca

Abstract—Recently, several car insurance companies got interested in classifying the behavior of drivers. Usage-based insurance (UBI), such as Pay-How-you-Drive (PHYD) scheme, is an innovative idea in which the insurance premium changes based on the driving behavior. This behavior is usually evaluated in terms of vehicle-related variables such as distance, speed, and acceleration to determine the expected risk profile for drivers. In this paper, an additional level of classification in the hierarchy of profiling is proposed. Using the 100-CAR naturalistic driving study (NDS) data set, five different Hidden Markov Models (HMMs) are trained to determine the fault responsibility of a Subject Vehicle (SV) in a crash or near-crash events. Two specific driving situations, which are conflicts with leading and following vehicles, are investigated in this study. Results show that these models can achieve a reasonable classification accuracy.

I. INTRODUCTION

A. Background

For a long time, car insurance policies have been viewed by many as unfair. That is because insurance premiums have been determined based on insufficient measures such as age, gender, etc. Consequently, some current insurance companies have adopted new adaptive premium schemes such as Pay-As-You-Drive (PAYD) and Pay-How-You-Drive (PHYD) schemes [1]. In these schemes, insurance premiums are adjusted based on the travel distance and the individual driver behavior, respectively. For this purpose, vehicle telematics manufacturers such as: Sky-meter, Onstar, Progressive, etc. have been collecting real-time data, such as vehicle's speed, acceleration, distance and fuel consumption data. This data is continuously measured using On-board diagnostics (OBD) tools (e.g. OBD II) or using smart-phone embedded sensors, and is then sent to a cloud for further processing. In PHYD schemes, driving behavior is conventionally characterized by the number of harsh braking, over speeding, and aggressive acceleration events. Based on these indicators, a risk probabilistic function is utilized to assign an individual risk profile for each driver [2]. Although this approach may provide a realistic risk estimation in some cases, it is too simplistic and can be unfair in other cases.

In fact, profiling drivers solely on such events without considering the driving situation in which they occurred can be misleading. That is because drivers vary in their responses towards driving conflicts according to their individual traits. For instance, some drivers show an extremely careful attitude even in un-risky events. This, in turn, causes them to do more frequent harsh braking events than other drivers. By deploying the current conventional PHYD schemes, these drivers will have unfair insurance premiums that are similar to careless and inattentive drivers. As a result, modified techniques that incorporate driving scene variables, which reflect the actual behavior of individual drivers in risky situations, are crucial.

B. Related Work

Driver profiling and classification has been an active area of research within the last few years. Reference [3] proposes a smart-phone application (i.e. Sensefleet) that is based on a fuzzy logic inference system. The system utilizes smart-phone's accelerometer data, gravity and magnetic sensory data fused with GPS data to detect hard acceleration, harsh braking, over speeding and aggressive steering events. Similarly, the work presented in [4] uses smart-phone's accelerometer, gyroscope, and magnetometer data to capture risky driving events such as sharp turning, aggressive acceleration and lane changing, and sudden braking. Signals are scaled using dynamic time wrapping (DTW) technique and maneuvers are then classified as risky or non-risky using a Bayesian classifier. MobiDriveScore application [5] utilizes both smart-phone as well as OBD II sensory data to determine risky events. Other works that are based on advanced discriminative and generative modeling approaches have been proposed. Reference [6] proposes two algorithms that are based on support vector machines (SVM) and HMM to predict the behavior of drivers at intersections. A large naturalistic data set is utilized for models training and evaluation. In [7], the decision making process at high-speed signalized intersections has been modeled using a two state HMM. Sathyanarayana et al. [8] proposes two HMM based modeling approaches (namely, bottom-to-top and top-

to-bottom) to identify different driving maneuvers. Oliver et al. utilized HMMs and their extension (Coupled HMMs) to identify seven different maneuvers [9]. Other works [10], [11] also use HMMs to identify risky maneuvers or behaviors.

Despite all of these efforts, models capable of capturing drivers fault contribution in creating risky events are still a void. To this end, our study is meant to be a step towards a comprehensive novel classification approach in which drivers are classified based on their behavior in different risky driving situations.

In this paper, an HMM-based modeling approach is deployed to determine the fault responsibility of a Subject Vehicle (SV) driver during a crash and near-crash events. Five unique HMMs, representing 5 classes of behaviors, are trained and validated using the 100-CAR NDS data set [12]. Three of these models are utilized to classify an SV driver when he/she is involved in conflicts with following vehicles, while the other two are used to classify conflicts with leading vehicles, in normal road and weather conditions. The rest of the paper is structured as follows: first, the driving fault determination problem is discussed in Section II. Data description and the pre-processing process is then detailed in section III. In section IV, the fault determination classification approach using HMM modeling is described. The experimental results are then depicted in section V. Finally, the conclusion is drawn in section VI.

II. PROBLEM STATEMENT

Consider a traffic environment in which traffic flow is in a single direction (i.e. one-direction or a divided roadway). The objective is to determine the fault responsibility of an SV driver during risky conflicts with following or leading vehicles (i.e. conflicts happening in the same lane). This is achieved via the analysis of the pre-incident maneuver that caused this risky event, and the SV re-action in response to that event. Figure 1 depicts an example of a conflict between an SV and a leading vehicle in a divided roadway.

By analyzing vehicle related signals (i.e. longitudinal and lateral acceleration) as well as range related signals (i.e. range, range rate etc.) different behavioral classes are assigned to the SV driver, depending on the conflict type. The classification process is detailed in section IV. This classification is in agreement with the Revised Regulations of Ontario (R.R.O.) 1990, Regulation 668, fault determination rules, under “*Rules For Automobiles Traveling In The Same Direction And Lane*” section [13].

The accurate determination of the SV’s behavior during risky driving conflicts lays the foundation for more advanced profiling techniques that can be utilized by insurance companies as well as fleet administrators.

III. DATA SET DESCRIPTION AND PRE-PROCESSING

A. 100-CAR NDS Data Set

In this study, the 100-CAR NDS data is used for model training and validation. The 100-CAR NDS project is a large-scale data collection project sponsored by the National High-

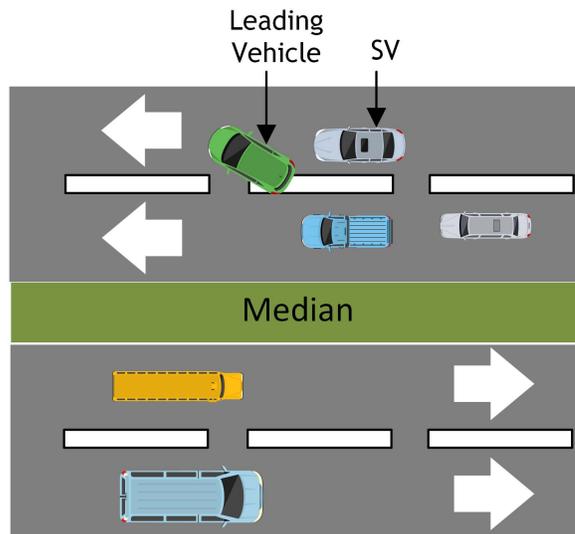


Fig. 1. A conflict with a leading vehicle in a divided roadway.

way Traffic Safety Administration (NHTSA) and the Virginia Department of Transportation (VDOT) [12]. In the 100-CAR NDS, 241 primary and secondary drivers were recruited over a period of 1 year to collect large-scale driving data. Recruited drivers used approximately 100 cars instrumented with a set of sensors including: forward and rearward radar sensors, OBD dongles, accelerometers, gyroscopes, five channels of digital video, and GPS. The sampling rate of the acquired data ranged from 1 Hz to 10 Hz. Data was recorded using electronic digital recorders (EDRs) resulting in a significant amount of driving data which is approximately 43,000 hours of data [12].

Data reductionists identified a total of 69 crash and 760 near crash events based on different types of triggering signals (e.g. acceleration or deceleration ≥ 0.5 g coupled with a time-to-collision (TTC) of 4 seconds or less). The detailed identification process could be found in [12]. Only 176 events that match the scope of this study have been used for model training and validation purposes. For each of these events, a file that contains a time-series data spanning 30 seconds before and 10 seconds after the event is available. Time-series data include 31 variables (e.g. speed, acceleration, etc.) that describe the SV behavior during each event. The detailed narrative of each event was extracted and documented using the installed digital video cameras. These narratives are used in this work as a reference for the purpose of SV behavioral classification and performance evaluation.

B. Data Preprocessing and Feature Selection

During the data collection process, sensors failed to capture values of some variables at some time instants, causing some gaps in the data. In this work, missing data are approximated using linear interpolation. The data is then re-sampled into a 1s time span to reduce the computational cost. This is achieved by taking the median value in a 1 second span of the initially sampled data. Only a few variables are initially selected as

candidate features for model training. They can be categorized into two types:

- 1) Variables that reflect the longitudinal and lateral movements of the SV. These variables are: Gas pedal position, speed (CAN-bus), speed (GPS), yaw rate, vehicle heading (GPS), longitudinal acceleration, lateral acceleration.
- 2) Variables that reflect the relative motion of the SV to leading and following vehicles. These variables are: Forward range and range rate, rearward range and range rate.

The forward range is the distance between the SV and the leading vehicle, while the rearward range is the distance between the SV and the following vehicle, at any point in time. They are expressed mathematically as:

$$R_{Forward}(t) = |x_i(t) - x_{i+1}(t)| \quad (1)$$

$$R_{Rearward}(t) = |x_i(t) - x_{i-1}(t)| \quad (2)$$

where $R_{Forward}(t)$ and $R_{Rearward}(t)$ are the forward and rearward ranges at time t , respectively, $x_i(t)$, $x_{i+1}(t)$, and $x_{i-1}(t)$ are the positions of the SV, leading vehicle, and following vehicle at time t , respectively. The range rate is the rate of change of the distance between the SV and following or leading vehicles. It posses a negative value when the distance decreases over time.

The statistical dependencies among “type 1” variables are obtained through the calculation of their correlation matrix. A fixed correlation coefficient threshold of a value ≥ 0.8 is adopted to indicate redundancy between two variables. Feature space is reduced by removing all redundant variables, resulting in only three “type 1” candidate variables which are: speed (CAN-bus), longitudinal acceleration, and lateral acceleration. Finally, the speed variable is removed for more simplification since it does not improve the classification accuracy based on the experimental results.

IV. FAULT DETERMINATION

The knowledge of longitudinal and lateral behavior as well as the relative position of the SV prior and during a risky event, can give an insight on driver’s contribution in creating such event. This leads to a more accurate and fair classification that takes into account the whole driving scene. The classification process depends on the driving conflict type in which the SV is involved. In general, drivers are classified to belong to one of the following four classes: *Faulty*, *Non-faulty*, *Skilled*, and *Non-skilled*. Each of these classes represents a set of unique and distinguishable observation sets. Faulty and non-faulty classification modeling is done in accordance with the insurance act rules in the case of crash occurrence. The other two classes represent a finer classification based on the driver’s reaction during the event. The latter two classes, although they are conventionally considered part of the non-faulty category, they have unique characteristics that single them out from the non-faulty class. In the rest of this section, we discuss two of the most common driving conflicts which are: conflicts

between the SV and leading/following vehicles. The devised models are built under three assumptions: 1) Normal weather and road conditions. 2) vehicles are in a one way or divided roadways. 3) Vehicles are not at an intersection.

A. Conflicts with leading vehicles (type 1)

Consider a situation where an SV is involved in a crash/near crash event with a leading vehicle. Table I summarizes the description of the different SV’s behavioral classes. The 100-CAR data set does not contain events that represent the non-skilled and non-faulty classes for this conflict type and under the aforementioned assumptions. The SV driver is classified as non-faulty when another driver cuts in too close in front of him/her, making it impossible to avoid crashing. On the other hand, the SV driver is non-skilled when an avoidable crash occurs due to the poor judgment or the inattentiveness of the SV driver. Hence, the SV in this class is still considered as non-faulty since he did not initiate the faulty maneuver. The other two classes will be explained using two sets of real observations obtained from the data set.

TABLE I
BEHAVIORAL CLASSES OF AN SV DRIVER INVOLVED IN A CONFLICT WITH A LEADING VEHICLE.

Class	Description
Faulty	Being involved in a crash/near-crash event with a leading vehicle due to the SV inattentiveness/aggressive behavior.
Non-skilled	Being involved in an avoidable crash event where the leading vehicle is faulty. The SV could safely take a corrective reaction to avoid the crash.
Non-faulty	Being involved in an unavoidable crash event where the leading vehicle is faulty.
Skilled	Avoiding a crash with a faulty leading vehicle by decelerating or changing lanes.

1) *Faulty class*: An SV is considered faulty in a conflict with a leading vehicle when he/she is not keeping enough forward distance prior to the conflict. Figure 2 depicts four time-series observations during a near-crash event. The observations from top to bottom are: longitudinal acceleration, lateral acceleration, forward range and range rate. Figure 2 shows that the SV has kept a constant speed (i.e. zero longitudinal acceleration) even with a decreasing forward range and range rate (i.e. the SV is aggressively speeding). The driver had to perform a harsh braking at the last moment (i.e. at time = 6s on the Figure) to avoid rear-ending the leading vehicle.

2) *Skilled class*: The skilled behavioral class comprises SVs who react skillfully to avoid a crash with a leading vehicle that cuts in too close in front of him/ her (i.e. faulty lane change maneuver). Figure 3 depicts the time-series data of a skilled driver during a risky conflict with a leading vehicle. Observations show a steep change in the forward range of the SV at time ≈ 5 s. The range is almost constant when it suddenly and significantly decreased. This reflects the unsafe lane change maneuver performed by a leading vehicle. The

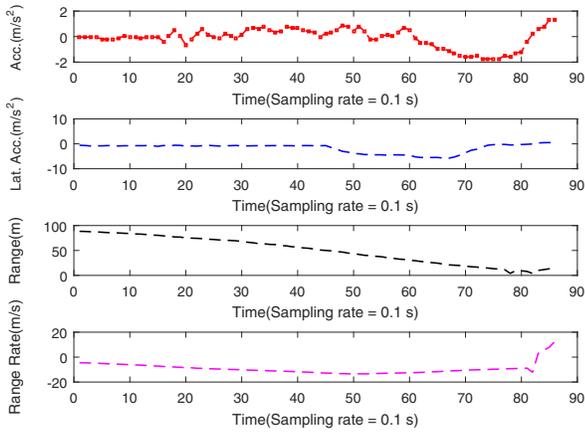


Fig. 2. A set of observations showing a faulty SV during a conflict with leading vehicle

sudden change in the lateral acceleration signal shows that the SV has avoided the crash by immediately making a left lane change maneuver followed by a smooth longitudinal braking.

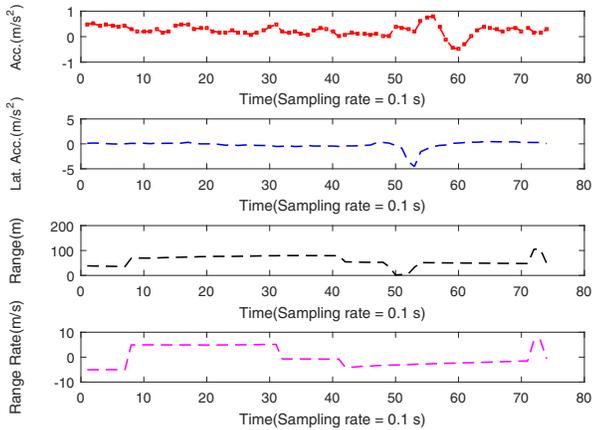


Fig. 3. A set of observations showing a skilled SV during a conflict with leading vehicle

B. Conflicts with following vehicles (type 2)

In table II, a brief description of different behavioral classes of an SV involved in a risky conflict with a following vehicle is presented. Based on the event narrative data dictionary of the 100-CAR data set, and the description of each class, the skilled behavioral class is missing in the available data. An SV driver is considered skilled in this type of conflicts when he/she reacts to an inattentive or aggressive following vehicle by safely accelerating or changing lanes. The other three classes are detailed in the following sections.

1) *Faulty class*: An SV is classified as faulty in this category of conflicts when he/she changes lanes without keeping a safe gap with the following vehicle in the new lane. Figure

TABLE II
BEHAVIORAL CLASSES OF AN SV DRIVER INVOLVED IN A CONFLICT WITH A FOLLOWING VEHICLE.

Class	Description
Faulty	Cutting too close in front of a following vehicle causing an unsafe rearward range or a negative range rate high in magnitude.
Non-skilled	Braking too hard in front of a following vehicle causing an unsafe rearward range or a negative range rate high in magnitude.
Non-faulty	Braking smoothly. Unsafe rearward range due to the reckless behavior of the following vehicle.
Skilled	Avoiding a crash with a following vehicle by accelerating or changing lanes.

4 shows an example of a set of observations that reflects this behavior. As can be deduced from this figure, the change in the lateral acceleration at time ≈ 4 s is followed by a steep change in the rearward range (i.e. from $R_{Rearward} \approx 43$ m to $R_{Rearward} \approx 5$ m). This is interpreted as a faulty lane change maneuver.

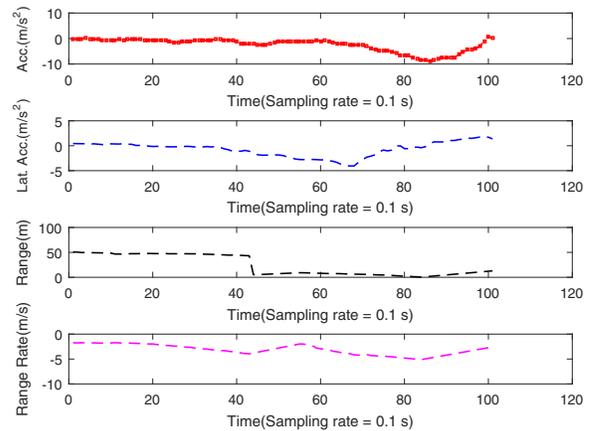


Fig. 4. A set of observations showing a faulty SV during a conflict with a following vehicle

2) *Un-skilled class*: In this class, the SV is in the same lane of the following vehicle when he/she makes an unnecessary harsh braking, causing a crash or near crash event with the following vehicle. Although the following vehicle is still classified as faulty, since it has to keep a safe rearward range, the conflict could be avoided if the SV did not make this unnecessary action. Figure 5 depicts a set of observations showing an unskilled behavior. The SV in this example performed an aggressive deceleration event at time ≈ 3 s. This has led the following vehicle to nearly hit the SV in the rear as can be shown from the rearward signal. In this incident, The forward range was big enough that the SV did not have to take such aggressive action.

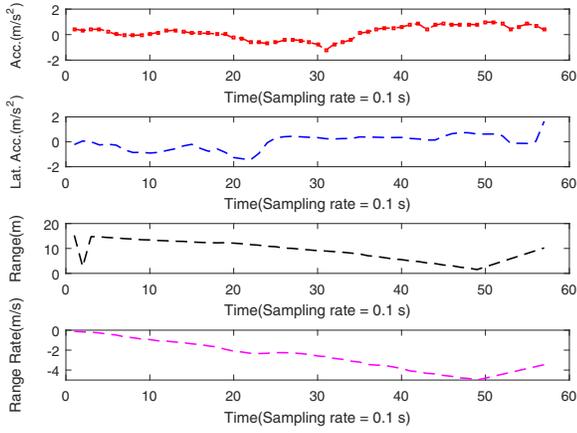


Fig. 5. A set of observations showing an non-skilled SV during a conflict with a following vehicle

3) *Non-faulty class*: In this conflict type, despite of the smooth behavior of the SV, the inattentiveness or aggressiveness of the following vehicle causes a risky event. In figure 6, an SV is involved in a near-crash event due to the inattentiveness of the following vehicle. As can be noticed, although the SV is keeping a constant speed, his rearward range started to decrease gradually at time ≈ 2.5 s. A near crash event occurred at time ≈ 5 s when the SV smoothly decelerated at time ≈ 4.5 s.

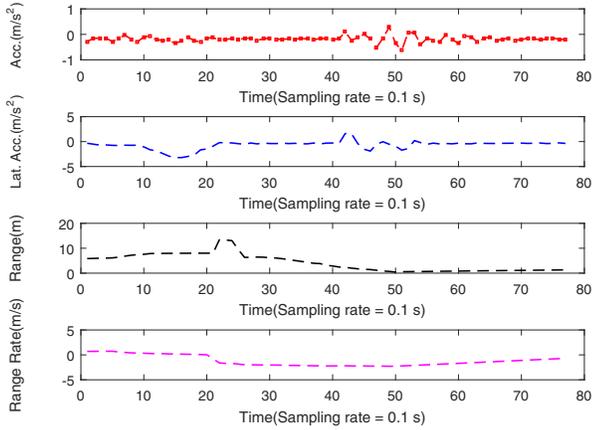


Fig. 6. A set of observations showing a non-faulty SV during a conflict with a following vehicle

C. HMM classification models

First order time homogeneous discrete HMMs are utilized to classify drivers under the two different conflict types. Hidden Markov modeling is one of the best modeling approaches for modeling time-series systems [14]. Unlike conventional Markov models, where each state corresponds to a fully observed physical event, HMM observations are stochastically

related to the hidden states. An HMM is formally defined by the five tuple $\Omega(N, M, A, B, \pi)$, where N is the number of states, M is the number of possible observations per state, A is the transition matrix, B is the emission matrix, and π is the initial state distribution array. Transitions between states are annotated with probabilities that compose the transition matrix (i.e. $P(S_i|S_j)$), whereas the emission matrix includes the emission probabilities of observation symbols given the presence at a specific state (i.e. $P(O(m)|S_i)$). The initial state distribution matrix contains the initial probabilities of each state (i.e. $P(S_i)$). For convenience, HMMs are usually described using the following simplified notation: $\lambda = (A, B, \pi)$.

In our problem, the number of hidden states represents the possible driving behavioral modes, and the observations are the six time-series signals previously discussed in section III. Each of the six observation sequences is quantized into 6 possible quantization levels which take values from the set $Q = \{1, 2, 3, 4, 5, 6\}$. For each risky event, the SV driver is characterized by these sequences, which are mapped into a 1-dimensional emission array. Some of the emission arrays are used for models training. Consequently, the Baum Welch algorithm, aka forward-backward algorithm, is used to solve the HMM learning problem. Baum-Welch algorithm is an iterative update procedure in which A , B , and π are re-estimated in each iteration to maximize the likelihood of a given observation sequence (i.e. $\max_{A,B,\pi} P(O)$). This is done by utilizing the following two update equations:

$$\xi_t(i, j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}, \quad (3)$$

and

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j), \quad (4)$$

where $\alpha_t(i)$ and $\beta_{t+1}(j)$ are the forward and backward HMM variables for states i and j , at time instants t and $t + 1$, respectively, a_{ij} is the transition probability from state i to state j , and $b_j(O_{t+1})$ is the emission probability of the observed symbol at state j and time $t + 1$. Readers are referred to [14] for a detailed explanation of this algorithm.

Five unique HMMs, representing the five behavioral classes discussed in the previous section, are trained using Baum Welch algorithm. L_k observation sequences are used to train each model, where k refers to the model index. Baum Welch algorithm does not guarantee the convergence to a global maximum. As a result, many initializations of the transition and emission matrices are tested to improve the converged values over different trials. A tolerance value of $1e^{-4}$, which represents the difference in the forward log probabilities between two successive iterations, is utilized as a convergence criteria. Model evaluation is performed using Z_k validation sequences. The evaluation process utilizes the well known Forward algorithm. For type 1 conflicts, two probabilities are computed for each sequence: $P(Z_k|\lambda_{Faulty})$, and $P(Z_k|\lambda_{Skilled})$. Similarly, three forward probabilities are computed for each sequence resulted from a conflict of type 2:

$P(Z_k|\lambda_{Faulty})$, $P(Z_k|\lambda_{Non-skilled})$, and $P(Z_k|\lambda_{Non-faulty})$. In each of these types, The SV is assigned to the class with the highest forward probability.

V. RESULTS

HMM algorithm is implemented using MATLAB R2015b statistics toolbox. A total of 176 risky sequences are used for models training and validation. 142 of these sequences correspond to risky events of type 1. 25% of these sequences are used to train the two aforementioned models. On the other hand, only 34 events are available for the case of conflicts with a following vehicle. Consequently, 50% of the data is utilized for models training in this conflict type. The original sequences are cropped to include only 2 s before the occurrence of the risky event. This has resulted in the best experimental classification results. As previously mentioned, the data is down-sampled to 1 Hz to minimize the computational cost. The overall accuracy of driver classification under each conflict

TABLE III
CONFUSION MATRIX FOR CLASSIFICATION UNDER TYPE 1 CONFLICTS

		Predicted	
		Faulty	Skilled
Actual	Faulty	77	18
	Skilled	3	9

type is calculated using the following formula:

$$Accuracy = \frac{T_P}{T_P + F_P} \quad (5)$$

where T_P and F_P are the numbers of true positive and false positive predictions for the considered class, respectively. The overall accuracy of type 1 and type 2 conflicts is 80.37% and 72.2%, respectively. Tables III and IV depict the confusion matrices of the different classes under the two conflict types. It is evident that using a larger data set in type 1 conflicts has led to better performance results.

TABLE IV
CONFUSION MATRIX FOR CLASSIFICATION UNDER TYPE 2 CONFLICTS

		Predicted		
		Faulty	Unskilled	Non faulty
Actual	Faulty	6	0	1
	Unskilled	0	7	2
	Non faulty	0	2	0

VI. CONCLUSION

This paper introduced a novel HMM-based classification approach for determining the fault responsibility of drivers involved in risky driving conflicts. The paper focused on two types of driving conflicts which are conflicts with leading and following vehicles, under normal weather and road conditions. A total of 52 training sequences were used to train five unique HMMs that represent five distinguishable behaviors. The models were successfully validated using 124 evaluation

sequences. Model training and evaluation were performed using Baum welch and Forward algorithms, respectively. An overall classification accuracy of 82.37 % is achieved for conflicts of type 1, whereas an accuracy of 72.2 % is achieved for conflicts of type 2. These promising results can be improved by using a larger data set with more training and validation sequences.

ACKNOWLEDGEMENTS

This work is funded by project no. 31R014-Research Center RTTSRC-4-2013 provided by the Roadway Transportation & Traffic Safety Research Center, United Arab Emirates University. This research is also supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant number: STPGP 479248.

REFERENCES

- [1] D. I. Tselentis, G. Yannis, and E. I. Vlahogianni, "Innovative motor insurance schemes: A review of current practices and emerging challenges", *Accident Analysis and Prevention*, vol. 98, pp. 139-148, Jan. 2017.
- [2] P. Handel, I. Skog, J. Wahlstrom, F. Bonawiede, R. Welch, J. Ohlsson, and M. Ohlsson, "Insurance telematics: opportunities and challenges with the smartphone solution", *IEEE Intell. Transp. Syst. Magazine*, vol. 6, pp. 57-70, Oct. 2014.
- [3] G. Castignani, T. Derrmann, R. Frank, and T. Engel, "Driver behavior profiling using smartphones: A low-cost platform for driver monitoring", *IEEE Intell. Transp. Magazine*, vol. 7, pp. 91-102, Jan. 2015.
- [4] H. Eren, S. Makinist, E. Akin, and A. Yilmaz, "Estimating driving behavior by a smartphone", in *Intelligent Vehicles Symposium (IV)*, 2012 IEEE, 2012, pp. 234-239.
- [5] T. Chakravarty, A. Ghose, C. Bhaumik, and A. Chowdhury, "MobiDriveScore - A system for mobile sensor based driving analysis: A risk assessment model for improving one's driving", in *Sensing Technology (ICST)*, 2013 Seventh Intl. Conf. on 2013, pp. 338-344.
- [6] G. S. Aoude, V. R. Desaraju, L. H. Stephens, and J. P. How, "Driver behavior classification at intersections and validation on large naturalistic data set", *IEEE Trans. Intell. Transp. Syst.*, vol. 13, pp. 724-736, June 2012.
- [7] K. Tang, S. Zhu, Y. Xu, and F. Wang, "Modeling drivers' dynamic decision-making behavior during the phase transition period: An analytical approach based on Hidden Markov model theory", *IEEE Trans. Intell. Transp. Syst.*, vol. 17, pp. 206-214, Jan. 2016.
- [8] A. Sathyanarayana, P. Boyraz, and J.H. Hansen "Driver behavior analysis and route recognition by Hidden Markov models", in *Vehicular Electronics and Safety*, 2008. ICVES 2008. IEEE Intl. Conf. on, 2008, pp. 276-281.
- [9] N. Oliver, and A. P. Pentland "Driver behavior recognition and prediction in a SmartCar", in *Proc. SPIE Aerosense 2000, Enhanced and Synthetic Vision*, vol. 4023, pp. 280-290, Apr. 2000.
- [10] S. Daptardar, V. Lakshminarayanan, S. Reddy, S. Nair, S. Sahoo, and P. Sinha, "Hidden Markov model based driving event detection and driver profiling from mobile inertial sensor data", in *Sensors*, 2015 IEEE, Nov. 2015.
- [11] P. Boyraz, M. Acar, and D. Kerr "Signal modelling and Hidden Markov models for driving manoeuvre and driver fault diagnosis in an urban road scenario", in *Intell. Vehicles Symp.*, 2007 IEEE, pp. 987-992, June 2007.
- [12] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, and J. Jermeland, "The 100-car naturalistic driving study, Phase II - results of the 100-car field experiment", Washington, DC:NHTSA.
- [13] R.R.O. 1990, Reg. 668: FAULT DETERMINATION RULES. [Online]. Available: <https://www.ontario.ca/laws/regulation/900668/>, accessed: Apr. 17, 2017.
- [14] L. R. Rabiner "A tutorial on hidden Markov models and selected applications in speech recognition", *Proc. IEEE*, vol. 77, no. 2, pp. 257-285, Feb. 1989.