Learning-based driving events classification

Claire D’Agostino¹ Alexandre Saidi² Gilles Scouarnec³ and Liming Chen⁴

Abstract—Drivers typically depict different behavior with respect to various driving events. The modeling of their behavior enables an accurate estimation of fuel consumption during the truck design process and is also helpful for ADAS in order to give relevant advices. In this paper, we propose a learning-based approach to the automatic recognition of driving events, e.g., roundabouts or stops, which impact the driver behavior. We first synthesize and categorize meaningful driving events and then study a set of features potentially sensitive to the driver behavior. These features were experimented on real truck driver data using two machine-learning techniques, i.e., decision tree and linear logic regression, to evaluate their relevance and ability to recognize driving events.

I. INTRODUCTION

A. Motivation

It is a common known fact that driver style impacts fuel consumption. Within the context of constantly depletion of oil, it is increasingly important to model real driving styles during the vehicle design process, e.g., truck; or to recognize driving style on short events for personalized ADAS (advanced driver assistance systems). Indeed, different driving styles can lead to significant variations in fuel consumption. A test carried out by Maincent in 2005 [1] with 33 drivers showed a gap from -11% to +14% L/100km between different drivers on the same cycle with the same heavy vehicle and similar traffic conditions. This study thus reveals that a change of driver behavior can bring 10% of fuel economy (depending also on road type) which is significant for freight companies, especially when this fuel economy does not imply a lower average speed. This last property is very important. Indeed, while we speak about eco-driving for personal cars, this term cannot apply to trucks. Eco-driving includes several rules such as anticipate traffic flow or shift up early, and thus it can result on a lowest average speed. As for truck drivers, they cannot reduce their average speed because of delivery imperatives due to their job. In the latter case, we rather refer to rational driving, term that Maincent [1] defines as: the full exploitation of the potential of the vehicles to drive more fuel-efficiently while respecting delivery delays. Vehicle driver typically depicts different behaviors depending on driving events that he has to face on the road. For instance a driver will not react similarly to a traffic light or a toll. This difference of behavior impacts on fuel consumption. Therefore, before being able to model drivers behavior the first stage is to recognize what are the different significant driving events. To reach this goal we propose in this paper a study on the machine-based recognition of significant road events using real truck driver data.

B. State of the art

Since driving is a daily task, everyone can understand that drivers behave differently depending on various parameters: vehicle, traffic, mood of the day... There are also many ways to study these different behaviors: one can focus on fuel consumption, another on average speed, or mental load, drowsiness, etc. Therefore, relevant parameters to characterize the driver behavior are application dependent. For instance, from the viewpoint of fuel economy, one will have a closer look on throttle pedal and engine torque whereas to analyze mental load parameters like blood pressure or visual activity will have more impact on classification. Concerning the driving data from which various parameters or features are extracted to analyze, they can be of three different length of patterns described in [2]:

- Large-scale driving patterns: long term data which allows a large analysis. These patterns often determine driving context but they can also be used to classify drivers (e.g. gas pedal signal of 5 minutes used in [3]).
- Small-scale driving patterns: these data are collected during drivers manoeuvres (changing lane, braking for a roundabout) therefore their length is around a few seconds.
- Real-time data: for example eyes or head position used to classify in-vehicle operations. The order of magnitude of the duration of this kind of data is about some milliseconds.

Ericsson [4] explains that driving patterns can be described by many parameters, average speed being the most common. In [4], [5], [6] and [7] the driving parameters presented in table I were introduced. All these features are generally used for the analysis of large-scale driving patterns. They can be used as such using their raw values, or be further pre-processed. For example in [3], velocity, following distance, gas pedal position and brake pedal position signals were modeled with Gaussian Mixtures Models as well as cepstral analysis before classification. Driving data is also used to analyze driving situations such as road type and traffic. All these parameters are relevant ones, but to our knowledge, they were suggested generally for large-scale patterns. In this work we intend to characterize driving events which are actually small-scale driving patterns. As a result, we will
experiment a selection of the previous features to check their relevancy in addition to some new ones that we are proposing as well.

### TABLE I
DRIVING FEATURES FROM LITTERTATURE

<table>
<thead>
<tr>
<th>Group</th>
<th>Driving parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>Average speed</td>
</tr>
<tr>
<td></td>
<td>Average driving speed (w/o) stops</td>
</tr>
<tr>
<td></td>
<td>Relative distribution of speed</td>
</tr>
<tr>
<td></td>
<td>Joint distribution of speed</td>
</tr>
<tr>
<td></td>
<td>Frequency of oscillations of the speed curve per 100s</td>
</tr>
<tr>
<td>Acceleration</td>
<td>Relative distribution of acceleration and deceleration</td>
</tr>
<tr>
<td></td>
<td>Joint distribution of acceleration and deceleration</td>
</tr>
<tr>
<td></td>
<td>Average acceleration (when (a &gt; 0.1 \text{ m/s}^2))</td>
</tr>
<tr>
<td></td>
<td>Average deceleration (when (d &lt; -0.1 \text{ m/s}^2))</td>
</tr>
<tr>
<td></td>
<td>Acceleration standard deviation</td>
</tr>
<tr>
<td></td>
<td>Average number of acceleration-deceleration changes within driving period.</td>
</tr>
<tr>
<td>Noticeable sub-parts</td>
<td>Proportion of standstill time (\text{(speed} &lt; 2\text{km/h})) (</td>
</tr>
<tr>
<td></td>
<td>Proportion of acceleration time (a &gt; 0.1 \text{m/s}))</td>
</tr>
<tr>
<td></td>
<td>Proportion of time at constant speed (</td>
</tr>
<tr>
<td></td>
<td>Proportion of deceleration time (a &lt; -0.1 \text{m/s}))</td>
</tr>
<tr>
<td></td>
<td>Percentage of time when acceleration exceeds 2.5m/s</td>
</tr>
<tr>
<td></td>
<td>Percentage of time speed &lt; 2km/h</td>
</tr>
<tr>
<td></td>
<td>Average duration of running periods.</td>
</tr>
<tr>
<td>Independant</td>
<td>Duration</td>
</tr>
<tr>
<td></td>
<td>Number of stops per kilometer.</td>
</tr>
<tr>
<td>Energetic-power analysis</td>
<td>Positive kinetic energy</td>
</tr>
<tr>
<td></td>
<td>Relative positive acceleration</td>
</tr>
<tr>
<td></td>
<td>Variables representing surrogates for inertial power ((\text{acceleration}^<em>\text{speed})) and drag power ((\text{acceleration}^</em>\text{speed}))</td>
</tr>
<tr>
<td></td>
<td>Percentage of time when speed*acceleration is superior to 3.6 \text{ m/s}^3</td>
</tr>
<tr>
<td>Engine</td>
<td>Others parameters concerning engine speed and gear shifting such as idle period</td>
</tr>
</tbody>
</table>

C. Contribution

The goal here is to recognize meaningful driving events based on small-scale driving patterns. For this purpose, we suggest to define first the significant driving events which impact on driver behavior. Then we introduce a set of features potentially relevant to describe these events. Finally we make use of two machine learning methods, namely decision trees and linear logistic regression models to classify driving events. The contributions of the paper are threefold:

- Identification and categorization of road events impacting driver behavior;
- Proposition and analysis of a set of features relevant to small-scale driving patterns while sensitive to road events;
- The use of two machine learning techniques, namely decision trees and linear logistic regression, for automatic recognition of road events from small-scale driving patterns.

The rest of the paper is organized as follows. Section II describes the data collection. Section III identifies the events which will be considered and their definition. Section IV focuses on features selection and section V introduces methods of classification. Section VI presents the experimental results using real data. Finally, section VII concludes this paper with discussion and future work.

II. DATA COLLECTION

A. First data acquisition

First data were collected with a Volvo vehicle (Medium Duty vehicle, distribution usage, see figure 1) with robotized gearbox, equipped with two sets of engine retarders (exhaust and compression) and loaded to 13.7 tons. The driver was asked to drive normally but without uses of kick-down. Data were recorded with a GPS-CAN, a data logger and a video recorder (not during all the time). This first acquisition campaign had been held in 2011 and it contains almost 19 hours of driving equivalent to 952km with the same driver on the same vehicle on 6 different paths around Lyon, France.

![Fig. 1. Vehicle used for data collection](image)

B. Validation data acquisition

In order to validate the work done with the first data, a new acquisition campaign has been held in 2012. It used the same vehicle loaded to 13.8 tons. Data were recorded with a GPS-CAN, a data logger along with a synchronized video recorded during all the campaign time. The same previous driver drove on two representative cycles (one long haul usage, one distribution usage) still around Lyon, France. These two cycles have been chosen depending on their events distribution: they have a similar percentage of each class of event than the whole initial database. It represents 4 hours of driving and 450km.

III. DRIVING EVENTS IDENTIFICATION AND DEFINITION

Instead of studying the global driving pattern for each cycle, we focused on small events to be able to model driving behavior for each event. Still keeping in mind that the final goal is to model a driver, events are defined depending on actuators use: brake, retarders, throttle pedal and none of these (coasting). For now, only brake and retarder actuators have been considered, together with the adjacent coasting periods. Throttle events will be considered later. A new event is defined each time service brake and/or retarders are activated. During the data acquisition we observed 16 causes which could provoke the use of brake and/or retarders: access ramp (i.e. braking for rejoining traffic), give way, speed bumps, red traffic light, truck size constraint (e.g. narrow road), low speed (i.e. start-stop events occurring in a traffic situation where speed does not exceed 10km/h), lower speed...
limit, toll, pedestrian crossing, keeping constant speed in downhill (i.e. because of the kinetic energy of the vehicle, the driver must brake in steep slopes in order to not excessively overspeed), roundabout, exit ramp, stop sign, traffic (e.g. traffic jam still on movement, unexpected car fitting into the traffic), roadwork and curve. Access ramp, roadwork and pedestrian crossing occurs less than 1%; it presents an overfitting risk therefore these occurrences were switched to constant speed downhill, truck size constraint or low speed according to the video recording. "Driver behavior" may have several meanings depending on the objective of the study. For this work, driver behavior represents how the driver controls his vehicle to adapt to the driving events: why and how he uses the brakes, the retarders. Thus at the beginning of the work on data, there were 913 supervised events and 13 selected classes. Since the focus was on driver behavior, the classes were identified with respect to driver actions and bound to the features presented in the following part.

### IV. Driving events sensitive features

As seen in the State of the art paragraph, main features concern speed and acceleration. All the features retained for classification are shown in table II. This table is the list of all the features which have been tested for classification. Results concerning what are the most sensitive features are in the experiments part.

#### V. Driving events classification

##### A. Decision tree

This method has been chosen because it keeps the physical sense of attributes which is highly necessary for this work. Indeed it was a priority for us to fully understand why some instances are classified in one class instead of another. Besides it allows to easily seeing what features are the most important thanks to the tree. We used C4.5 algorithm in WEKA [8] data mining software to classify the events. It is called J48 and is based on [9]. It builds a decision tree using the information entropy. That means that at each node the most discriminant attribute is determined depending on information gain, weighted by the number of instances in each class (gain ratio). The gain \( G(S,A) \) for attribute \( A \) of the set \( S \) is:

\[
G(S,A) = E(S) - \sum_{i=1}^{m} f_{S}(A_i) \times E(S_{A_i})
\] (1)

where \( m \) is the number of different values of \( A \) in \( S \), \( f_{S}(A_i) \) is the frequency of the items possessing \( A_i \) as value for \( A \) in \( S \), \( A_i \) is the \( i^{th} \) possible value of \( A \), \( S_{A_i} \) is a subset of \( S \) containing all items where the value of \( A \) is \( A_i \) and \( E(S) \) is the information entropy of the dataset \( S \):

\[
E(S) = -\sum_{j=1}^{n} f_{S}(j) \times \log_2 f_{S}(j)
\] (2)

where \( n \) is the number of different values of attributes in \( S \) and \( f_{S}(j) \) is the frequency of the value \( j \) in the set \( S \).

##### B. Linear logistic regression model classifier

In order to compare with another algorithm and to improve the results, we also classify with an algorithm of linear logistic regression. It still keeps the physical sense of the features, but it also builds a linear model with the features value. In WEKA software, it is called SimpleLogistic algorithm and it is based on [10] and [11]. It creates linear logistic regression models to obtain a probability which represents the posterior probability to belong to the class. The linear logistic regression associates a linear model with a probability function, and ensures that the probabilities sum to one and remain in [0,1]. [9] presents the form of the linear logistic regression model:

\[
P_c(G = j|X = x) = \frac{e^{F_j(x)}}{\sum_{k=1}^{J} e^{F_k(x)}}
\] (3)

where

\[
\sum_{k=1}^{J} F_k(x) = 0
\] (4)
\[ F_j(x) = \beta_j(T) \cdot x = \sum_{i=1}^{p} \beta_{ij} x_i + \beta_{j0} \quad (5) \]

where \( Pr(G= j|X = x) \) is the posterior class probability to belong to class \( j \) knowing the input vector of features values \( x = (x_1, \ldots, x_p) \), \( F_j(x) \) is the linear model with parameters \( \beta_j \). The simple logistic algorithm estimates these parameters based on the LogistBoost algorithm [12], which runs until convergence to find the maximum likelihood linear logistic model whom linear functions are fit using linear least square regression.

VI. EXPERIMENTS

A. Features selection

Algorithm computing gain information (see formula (1)) with respect to the class shows that the most basic features are the most discriminant: maximum speed during the event, minimum speed, gap between initial and final speed, road type, duration, distance and whether there is a stop or not. Then follow context attributes: gap distance with previous and next event, previous event, next event, and average speed within the last kilometer. Several classification tests have also been done with different subsets of attributes to illustrate this result. We tested growing subsets from 1 to all the attributes. Not all the subsets were tested because of computing time and also because some features quickly appeared to be not really discriminant. The results from select attributes algorithms oriented these tests which finally confirmed the previous result: speed attributes are the most sensitive features. Besides, the figure 2 shows that it is not the more attributes the better classification rate. Each point represents a test done with a subset of features (number of features on x-axis) and the corresponding classification result (y-axis). We observe that some features are really discriminant whereas others bring some noise. Finally the best results were obtained with these features: duration, distance, maximum speed, minimum speed, \( \Delta \) speed, speed standard deviation, road type, gap distance with previous and next event, average speed within the last kilometer and whether there is a stop.

B. Classification results

The goal of this classification work is to be able to recognize efficiently an event and to group similar events. This first work has been done on the initial database of 913 events. Unless indicated otherwise all the results presented below are obtained with C4.5 algorithm using 10-folds cross validation method. It partitions 10 samples of 91 or 92 instances. Then it repeats 10 times training on 9 samples and testing with the remaining sample. A quick overview of the database showed many inter-classes similarities and intra-classes dissimilarities. It was confirmed by clustering work and linear Fisher discriminant calculus:

\[ J_{1/2} = \frac{1}{n} \sum_{i=1}^{n} \frac{(\mu_{1,n} - \mu_{2,n})^2}{\sigma_{1,n}^2 + \sigma_{2,n}^2} \quad (6) \]

where \( J_{1/2} \) is the linear Fisher discriminant between class 1 and 2, \( n \) is the number of features, \( \mu_{C,n} \) is the average value of feature \( n \) in class \( C \) and \( \sigma_{C,n} \) is the standard deviation value of feature \( n \) in class \( C \). This criterion can be used for linear classification but in our case we consider it as a distance between two classes. Figure 3 illustrates distance between the classes, for example toll class is far away from all the other classes. It also shows how much close some other classes are: size constraint, lower speed limit, keep constant speed in downhill, roundabout and stop are all less than 1 to traffic class. With these similarity inter-classes and dissimilarity intra-classes, the first result of classification with C4.5 algorithm was around 49%, as we can see in table III. Since there are 13 classes, this result is however better than a random classification.

Watching closely speed graphs for class \( E: \text{size constraint} \), we see that 5 instances need to be switched to another class. Then it appears that for class \( F: \text{curve} \), high speed instances have similar shape to class \( H: \text{speed limit} \). These first changes improve classification rate to 63%, especially for \( H: \text{speed limit} \) class (+18%) and \( A: \text{roundabout class} \) (+10%), this class is now less confused with \( F: \text{curve class} \).
It then appears that exit ramp class seems to be difficult to classify and several trials have been done to mix this class with toll (during rolling, a toll and an exit ramp may be considered as similar) and speed limit (an exit ramp can be considered as a lower speed limit). However these tests were not conclusive since classification rate did not increase. Finally, still thinking that an exit ramp which concludes with a stop is similar to a toll, exit ramp instances with stop are considered as toll.

Then another thought about driving is that stopping for a traffic light can be similar to stopping for a 1: stop sign. This idea is confirmed by the part of the confusion matrix shown in the table IV. A confusion matrix is a table layout whose columns represent instances in predicted classes while rows represent instances in actual classes. The table show how the algorithm is confusing traffic light and stop sign. When these two classes are merged together, classification rate improves until 57%.

The following step concerns give-way class. Its classification rate is very low. Watching classification results, it occurs that give-way instances with stop are always considered as stop in C4.5 classification. Therefore these 5 instances are changed into stop class. Then the remaining give-way instances are merged with roundabout class because roundabouts are a variant of give-way, as checked with our internal road experts. These changes allow reducing the number of classes and improving classification result.

At this current step, classification results and confusion matrix shows that speed bump, size constraint and curve are often mixed. It is confirmed by observing speed curves of these events as we can see on figure 4.

These three classes can be grouped together in a big class named urban speed reduction. It groups all speed reduction events in urban area except those caused by traffic or roundabouts. It reduces the number of classes to 9 and classification rate is still improved from 58.5% to 63.5%.

![Fig. 4. Speed graphs for classes C, E and F](image-url)

The negative point of this classification is now the traffic class with a true positive rate around 20%. One possibility to improve this rate is to analyze this class separately and split it in different groups depending on some attributes. The first test was made on maximum speed feature: when superior to 50km/h, instances were labelled as lower speed limit and on the contrary they were labelled as urban speed reduction. The second test concerned road type attribute: when it is urban, instances were labelled as urban speed reduction, when it is main road, instances were labelled as lower speed limit and in the third case highway instances are labelled as traffic. With one or the other solution, global classification rate is finally around 74%. However second option is preferable since all rates are higher, whereas with the first option exit ramp true positive rate is 5% which is really low and not acceptable. Finally all these modifications involve a classification rate increase from around 50% to around 75% such as shown in table V. With linear logistic regression model algorithm, the results are still a little better: around 76%. During all the process we observed similar results between decision tree and linear logistic regression model algorithm, with the same trends. However decision tree seemed the better machine learning tool for this work, for the reason that it allows to understand more easily classification errors with its graphical representation. Moreover it gives rules classification as well.

C. Discussion

This classification of driving events was not an easy task because of the inter-classes similarity and intra-classes dissimilarity. That is why a final classification rate of 75% is a promising result given the nature of the data. It is understandable that some braking events for a roundabout

### TABLE III
TRUE POSITIVE RATE WITH INITIAL DATABASE

<table>
<thead>
<tr>
<th>Class</th>
<th>All</th>
<th>Roundabout</th>
<th>Traffic</th>
<th>Speed bump</th>
<th>Traffic light</th>
<th>Truck size constraint</th>
<th>Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP rate</td>
<td>49</td>
<td>60</td>
<td>30</td>
<td>43</td>
<td>77</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>Instances</td>
<td>913</td>
<td>146</td>
<td>127</td>
<td>60</td>
<td>129</td>
<td>47</td>
<td>95</td>
</tr>
</tbody>
</table>

### TABLE IV
PART OF THE CONFUSION MATRIX

<table>
<thead>
<tr>
<th>Stop</th>
<th>Traffic light</th>
<th>Classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>22</td>
<td>Stop</td>
</tr>
<tr>
<td>22</td>
<td>92</td>
<td>Traffic light</td>
</tr>
</tbody>
</table>
may have a similar approach to a size truck constraint event. That is why it cannot be expected to reach classification rate significantly higher than 75%. Moreover, since events are defined depending on brake and retarder actuation, they did not last all the same time. It was another difficulty to analyze data of different length, and we postponed other investigations such as spectral and cepstral analysis, as presented in [3], to focus in the first place on other features. Finally and hopefully, this work showed that basic features, especially maximum and minimum speed, were the most important. In order to validate this work, a second data acquisition has been realized one year after the first one (see paragraph II: Data collection). It represents 242 events with the same driver. Choosing the 913 initial events as training set and these new 242 events as testing set, the classification with C4.5 algorithm delivers a classification rate around 70% with selected attributes. This result confirms the work on the initial database and features.

VII. CONCLUSION AND FUTURE WORK

To the best of our knowledge analyzing and classifying short time driving events had not been done before. It can be used for different objectives, from modeling driver behavior in response to each event regarding their rational driving, to recognize events for personalized ADAS. However this work focused on the first step before reaching these objectives, i.e. identify the driving events. We suggested a selection and characterization of the events. The conclusion is that basic features (duration, distance, speed features, road type) added to context ones (distance gap with previous and next event) provide the best classification results. Despite the difficulty of classification due to raw nature of the data, this work finally reaches a satisfactory true positive rate of 75% on the 913 events. Moreover, the confusion matrix shows that the errors are mainly between similar classes such as urban speed reduction with roundabout. This explains why we consider the final classification as promising. This result has as well been validated with 242 new events. It is a really acceptable rate given the intra-classes dissimilarities and inter-classes similarities. It has been reached thanks to several handlings on the database, including data preparation and consolidation. Besides all the changes were realized in order to stay as close as possible to the real situation and to better merge similar events, always validated by our road experts. Regarding the future improvements, we noticed that temporal aspect may bring a different point of view therefore it could be interesting to do frequency or cepstral analysis. Another perspective concerns speed and acceleration data: since they are the most important features, we thought that it could be relevant to model them and add fitting parameters as features for classification.

However this learning-based driving events classification will be finally achieved when events concerning throttle will be also defined and classified. We can expect to reach at least similar classification rate with some similar attributes (speed in particular, which should still have a major role in classification) and maybe some new others features.

After this first stage focusing on driving events identification, the next step will consider rational driving with analysis of different driver behaviors during each class of events. Given the result of the driving events classification, event classes will be separated as a hierarchical classification. Within these classes different driving behaviors will be classified depending on fuel consumption and rational driving, and this work will allow to model and recognize driver behavior.

REFERENCES