October 2016

Identifying Individual Driver Behaviour Using In-Vehicle CAN-bus Signals of Pre-Turning Maneuvers

Mahboubeh Zardosht
The University of Western Ontario

Supervisor
Professor Michael A. Bauer
The University of Western Ontario

Graduate Program in Computer Science

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

Abstract

All drivers have their own driving style while performing different driving maneuvers. They vary in using vehicle's control devices such as the steering wheel, pedals, gears etc. In this thesis, we analyze driving behaviour in different timeframes prior to turns. We employ data obtained from actual driving behaviour in an urban environment collected from the CAN-Bus of an instrumented vehicle. Five CAN-Bus signals, vehicle speed, gas pedal pressure, brake pedal pressure, steering wheel angle, and acceleration, is collected for 5, 10, and 15 seconds of driving prior to each turn. We consider all turns for each driver as well as look specifically and right and left turns. We use cluster analysis to see if we can categorize drivers into possible groups of driving styles. In our first approach, we use hierarchical clustering on statistical features extracted from the signals. The results show that using this approach we can effectively cluster drivers into two groups, moderate and aggressive drivers. This pattern is also reflected in the analysis of right and left turns. Another approach makes use of the Dynamic Time Warping (DTW) technique to identify the distance between signals of each pair of drivers, and based on these distances, a cluster analysis using hierarchical clustering is performed as well. The results show high consistency in the membership within a cluster throughout different timeframes.

Keywords

Driving behaviour, Dynamic Time Warping (DTW), statistical feature extraction, Hierarchical Clustering Analysis (HCA).

Acknowledgments

I would like to express my greatest gratitude to my supervisor Prof. Michael A. Bauer for his tremendous support through this research, for his motivation and immense knowledge. Without his guidance and persistent help this thesis would not have been successful.

Besides, I would like to show my sincere thanks to the examining committee.

I consider it an honor to work with other members of our lab, especially, Dr. Steven Beauchemin, Cristian Ardelean, Jennifer Knull, Besat Zardosht, and Mohsen Zabihi.

I owe my deepest gratitude to my lovely husband, Hassan, for all the love, encouragement and continuous support he gave me throughout this research. This dissertation would not have been possible without his support.

Last but not least, my sincere appreciation to my parents, who always support me in every moment of my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

As the number of vehicles and road mileage increases, traffic safety has become one of the main issues for governments and manufacturers. Traffic accidents are one of the main reasons for injuries today. According to a study in 2015 for The National Highway Traffic Safety Administration (NHTSA)[1], in approximately 94% of the accidents examined, the major reason was driver's error. The result of this study, which was conducted based on data from the National Motor Vehicle Crash Causation Survey (NMVCCS) 2005-2007, is shown in Table 1-1.

Table 1-1. Critical reasons of crashes from 2005 to 2007 conducted by the National Motor Vehicle Crash Causation Survey (NMVCCS).

	Estimated	
Critical Reason	**Number of Crashes**	**Percentage***
Drivers	2,046,000	94% ±2.2%
Vehicles	44,000	2% ±0.7%
Environment	52,000	2% ±1.3%
Unknown	47,000	2% ±1.4%
Total	2,189,000	100%

*Percentage are based on unrounded estimated frequencies
(Data Source: NMVCCS 2005-2007)

In order to improve traffic safety and traffic efficiency, it is essential to try to understand the characteristics of driver behaviour and study the relationships between driving behaviour and traffic systems. Understanding driver behaviour can help us to build models of drivers which can be used to improve Advanced Driver Assistance Systems (ADASs), improve vehicle safety and privacy, and also help to detect risky driving styles.

[1] https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115

In this thesis we examine driving behaviour based on data collected from actual drivers in a controlled driving scenario.

1.1 Problem Statement

Driving is one of the most common yet highly complex tasks that individuals do. It involves multiple essential subtasks and it can be affected by many internal and external factors. Driving consists of a series of complex decisions and actions that a driver performs based on the current traffic environment. It is obvious that different drivers have different driving behaviour in the same traffic situation. Drivers differ in how hard they hit the pedals, in the way they turn the steering wheel, how they keep their eye on the road, how much distance they keep when following a car, etc. (Miyajima et al. 2007; Higgs and Abbas 2015). In fact, these differences are key factors in building individual drivers' driving behaviour. As a consequence, various driving behaviours need to be analyzed in order to personalize Intelligent Transportation System (ITS) applications for different drivers with different driving styles.

In addition to the ITS personalization, identifying different types of driver behaviour can also help improve traffic safety by identifying safe or unsafe driving styles (Chen, Pan, and Lu 2015), aggressive or normal drivers (Carmona et al. 2015; Johnson and Trivedi 2011), distracted or undistracted drivers (Choi et al. 2007), etc. Another application of driver behaviour analysis is in the area of security and privacy, for example, by looking at driver identification based on driving style (Enev et al. 2016).

Consequently, being aware of the differences between driving styles and behaviour, we can model intelligent Advance Driver Assistance Systems (ADAS), and improve the performance of each individual driver.

1.2 Research Approach

In this thesis, normal driving behaviour is analyzed in order to try to identify different driving styles. Our focus in this study is on the driver's behaviour before turns. Several important driving signals have been extracted from an automobile's Controller Area

Network (CAN-Bus)[1] data as driving indicators. Vehicle speed, gas pedal pressure, brake pedal pressure, steering wheel angle, and acceleration are the signals used in this study. Since the activities in preparing for a turn is a challenging and complex driving behaviour, and given that it has not been studied, we decided to focus on this activity. Our goal is to explore whether there are clusters of driver behaviour that can be identified from driving signals during small periods of time before turning maneuvers.

Our assumption is that that there are different styles of driving when approaching a turning maneuver. We apply two approaches to cluster the drivers based on their driving behaviour before turns. In the first approach we extract statistical features from the driving signals prior to each turn and look at clusters based on feature vectors. In the other approach, we use the Dynamic Time Warping technique to find the similarities between the extracted CAN-Bus signals in each driver prior to turns, and cluster based on those similarities.

1.3 Thesis Organization

The rest of this thesis is structured as follows. Chapter 2 consists of three different subsections covering the related literature. In the first section, background information on driving behaviour is discussed. The next section covers previous work on clustering driving behaviours, and the third section provides an overview of work making use of the Dynamic Time Warping technique. A description of the data we used for our study can be in Chapter 3. Chapter 4 consists of an explanation of our research approach. In Chapter 5, the results are presented and discussed. Finally, Chapter 6 presents our conclusions and discusses future work.

[1] CAN-Bus is a vehicle bus standard designed to allow microcontrollers and devices to communicate with each other.

Chapter 2

2 Literature Review

Since the 1950s, understanding and modeling various driving behaviours has always been the traffic scientist's issue of concern (Chandler, Herman, and Montroll 1958). In fact, analysis of driving behaviour is needed by many scientists and researchers. Intelligent vehicle designers need to understand driving behaviour in order to make driving assistance systems work properly in dynamic traffic situations. Autonomous vehicle designers need them to make driving driver-free. Traffic engineers need them to improve the safety and reliability of roads and related infrastructure.

Driving behaviour, like every other human-related task, is complicated and hard to analyze. Our research is focused on the analysis of driver data, specifically looking at driver behaviour based on in-vehicle signals using extracted statistical features and also using Dynamic Time Warping. In this Chapter, the research developments in the field of understanding driving behaviour, and the use of cluster analysis and Dynamic Time Warping in understanding driving behaviour are reviewed.

2.1 Understanding Driving Behaviour

Various literature has emphasized the need for a comprehensive method to understand basic normal driving behaviour and to distinguish between different driving styles. Drivers have different driving behaviours because there exist various different aspects of driving. They differ in how they use the pedals (gas, brake), how fast or slow they turn a steering wheel or how often they change speed. Several scholars have considered the role of these driving signals in characterizing driver behaviour.

Aarts and Van Schagen emphasized the role of vehicle speed in driving behaviour in both road and traffic safety (Aarts and van Schagen 2006). They discussed some of the research which studied the relation between vehicle speed and probability of crashes and indicated that high speed not only makes collisions more severe, but also increases the

risk of accident occurrence. The relation between vehicle speed and the number of crashes was also investigated in a study conducted by (Fildes, Rumbold, and Leening 1991) and some of their results are illustrated in Figure 2-1. This Figure shows the relation between speed and number of crashes in both urban and rural roads which indicate that the higher the speed, the larger the increase in the crash rate.

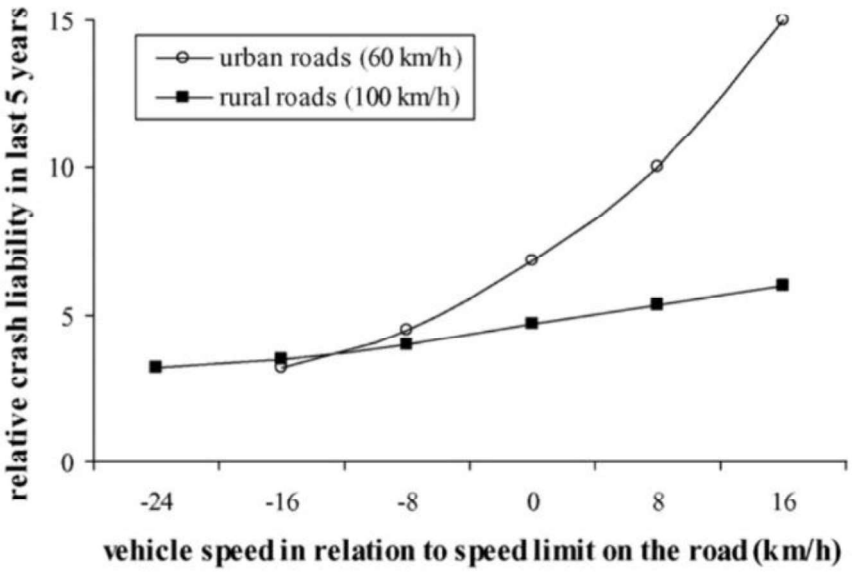

Figure 2-1. Relation between vehicle speed and number of crashes, based on self-reports in Australia, in urban and rural roads for last 5 years as found in the study of (Fildes, Rumbold, and Leening 1991).

(Miyajima et al. 2007) conducted a study on modeling driver behaviour in order to perform driver identification. In their study, they show that gas and brake pedal operation signals efficiently model individual driver differences. GMMs (Gaussian Mixture Model[1]) are used to model the spectral features of pedal signals and an identification rate of 76.8% is achieved, which suggests that each driver has a different pattern in pedal pressure. The spectral features are frequency based features which are

[1] A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities

obtained by converting the time based signal into the frequency domain using the Fourier Transform.

In another study, Ohta shows that driver behaviour differs among drivers considering the distance they keep when following another vehicle (Ohta 1993). He asked drivers to follow a car at various distances which are proper for them (e.g. most comfortable distance, minimum safe distance, etc.). Based on that he defined a temporal comfort zone for individuals following another vehicle; these were between 1.1 and 1.7 seconds. Below 1.1s he considered critical and more than 1.7s behind was considered uncomfortable because it was against the social norm. Consequently, based on these thresholds for a comfort zone, drivers' behaviour can be classified in three different groups considering how frequently a driver drove in a particular zone.

Chen, Pan, and Lu build a driving behaviour classification model to model safe and unsafe driving behaviours (Chen, Pan, and Lu 2015). In their study, vehicle operation data, including vehicle speed, engine RPM, throttle position, and calculated engine load, are collected via OBD (On-Board Device) interfaces and make use of AdaBoost algorithms for classification.

Choi et al. perform driver analysis using Hidden Markov Models (HMMs)[1] and Gaussian Mixture Model (GMM) on in-vehicle CAN-Bus signals such as steering wheel angle, brake status, acceleration, and vehicle speed (Choi et al. 2007). Three different classification tasks were conducted in this study, 1) action classification, 2) distraction detection, and 3) driver identification. Action classification consisted of categorizing long-term driving behaviours such as turning, lane changing, stopping, and constant driving. In distraction detection the goal was to detect if the driver was distracted by any secondary tasks such as working with cellphone, GPS, etc. Driver identification is concerned with driver classification based on driving behaviour characteristics. For driver identification they generate HMMs and GMM models based on driving signals collected

[1] A hidden Markov model (HMM) is a statistical Markov model in which the system is assumed to be a Markov process with hidden states and can be presented as the simplest dynamic Bayesian network.

during neutral and distracted driving periods. They classified the drivers using seventy percent of driving signals for training the models and the remaining was used to test the models. Figure 2-2 shows the accuracy of the best result for driver identification, for both 5 and 10 seconds segmented signals. This result is obtained by applying HMM using a 3 state model, and its best performance is when the number of Gaussian components is 7 (31.45% for 5 seconds and 29.16% for 10 seconds). They were able to identify drivers based on analysis of signals during distracted and neutral driving about 25% of the time; they did not evaluate their model on normal driving behaviour to study the differences between drivers in normal driving conditions.

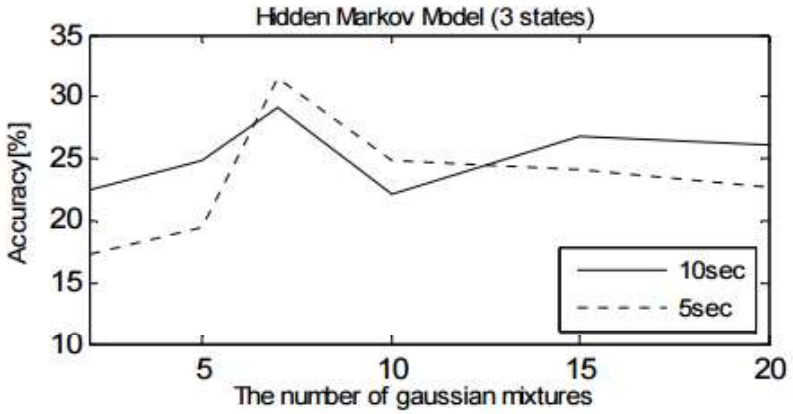

Figure 2-2 Accuracy of driver identification using HMM with 3 states for 5 and 10 seconds segmented signals (Choi et al. 2007).

2.2 Cluster Analysis Approaches

Several studies of driving analysis have been based on data mining in order to find a meaningful relation between data derived from different sources, as well as data from vehicle monitoring systems, driver behavioural characteristics, and road safety systems. Research utilizing cluster analysis is divided into two main groups, univariate clustering and multivariate clustering. In this section, we will review the literature on the research that utilizes cluster analysis to analyze drivers' behaviour.

In research performed by Kalsoom and Halim, individual driver behaviours are classified based on each driver's statistical driving features, such as, the ratio of indicators to turns,

the number of brake uses, the number horn uses, the average speed, the maximum speed and the gear (Kalsoom and Halim 2013). They apply K-means and hierarchical clustering on their experimental data in order to try to classify them to slow, normal and fast driving styles over the entire driving sequence. The data in this study was collected from a driving simulator and contained 5 minutes of data for each driver. The clustering was not based on the values of time-series driving signals, instead they used the number of operative devices, such as brake, gear, clutch, horn, left/right indicators. They also used the average and maximum of gears and average and maximum of speed in a 10 second time window. The results of the clustering analyses were inconclusive, with both Hierarchical Clustering Analysis (HCA) and K-means explored with different numbers of clusters. Clustering with HCA resulted in most of the data being grouped into a single cluster.

In a study by Wu et al., GPS data is collected from typical driving of commercial vehicles by professional drivers and then after using factor analysis the attributes which were related to the driving behavioural characteristics were extracted from the GPS data (Wu et al. 2016). Based on the GPS data, 8 speed and acceleration related features are extracted for each driver, and those were combined into four aggregated attributes: acceleration-deceleration, speeding-prone, acceleration, and deceleration. Using each of these attributes as indicators, four different cluster analyses of driver behaviour is done using hierarchical clustering. In each analysis, they identified five clusters that indicated how risky the driver behaviour is considering each attribute: minimal, slight, moderate, serious, or severe. Among the four cluster analyses were performed on 50 drivers, the results were not consistent across all attributes and drivers, but there were couple of drivers that showed a higher degree of risky driving behaviour for all the attributes. Again, they focused on trying to characterize the driving style of each driver over the entire driving period.

In both of these previous studies, the researchers focused on trying to use cluster analysis to categorize drivers across the entire driving sequence. This may be very challenging because over long periods driver behaviour may “average out” to look very similar. In

our work we have chosen to focus on a short driving period immediately before an identifiable event, i.e., a turn.

In 2013, Higgs and Abbas investigate the hypothesis that drivers have different driving behaviour in their daily driving tasks (Higgs and Abbas 2013). In order to investigate this assumption, three different truck drivers representing low-, medium-, and high-risk drivers performed 10 different car-following periods. They first divided the car-following period into different segments and then performed a cluster analysis on these segments. For clustering and optimization, they use different K-means techniques. Based on their results, each of three drivers shows a specific distribution of behaviours, however some of these behaviours are common between drivers but at different frequencies. For example, the behaviour of tailgating occurred in high frequency in the high-risk driver, low frequency in the medium-risk driver and did not occurr in the low-risk driver behaviour.

In another study done by Higgs and Abbas, a two-step algorithm was introduced to segment and cluster car-following behaviour based on eight state-action variables, in order to define driving pattern of drivers (Higgs and Abbas 2015). They defined driver behaviour as the way a driver responds to the current driving state (e.g., the vehicle speed, distance from the following vehicle, etc.) by performing a specific action (e.g., steer or push brake pedal). In this study, each car-following period was divided into similar driving state-action signals in the segmentation phase. Specific segments may repeat several times so the clustering phase tries to find and cluster the repeated segments into groups using K-Means as the clustering method. A representation of this two-step algorithm is shown in Figure 2-3 for a sample car-following period. In the middle graphs in this figure, segments of similar data have been formed. Then in the clustering step, these segments are processed and clustered as shown in the bottom graph. Since the clustering step is more constricting than the segmentation step, some adjacent segments are placed in the same cluster.

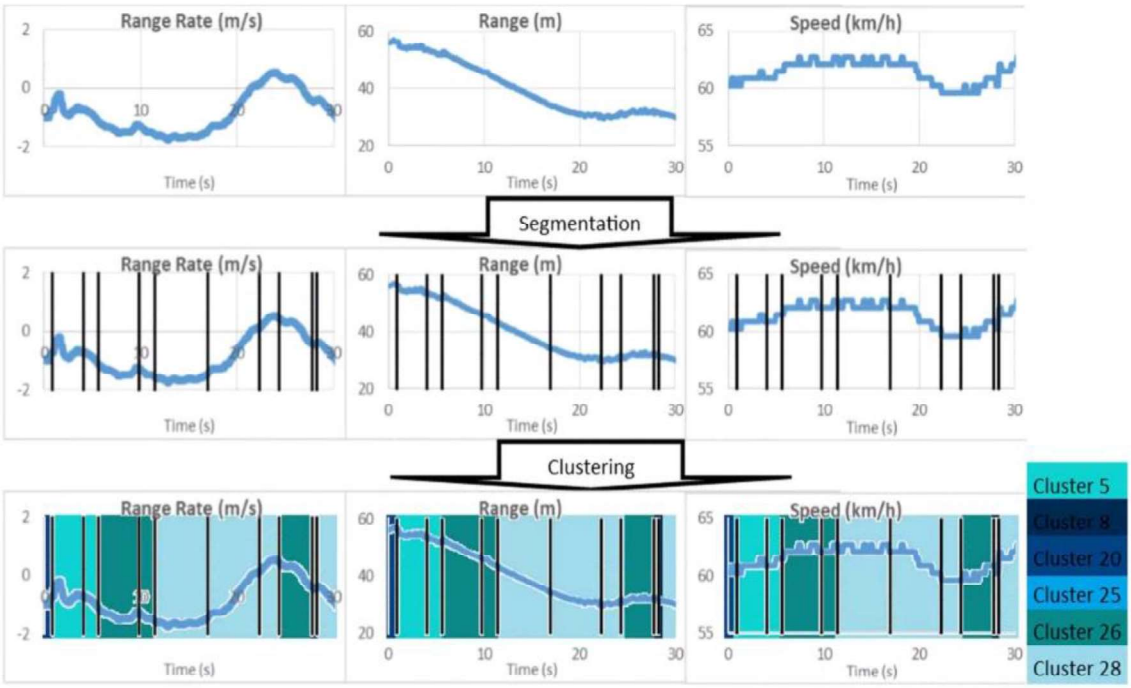

Figure 2-3. Two-step algorithm to segment and cluster driver car following behaviour (Higgs and Abbas 2015).

2.3 Dynamic Time Warping

Dynamic Time Warping (Berndt and Clifford 1994) is a time-series alignment algorithm, which attempts to align two sequences of features by warping the time axis to find an optimal match. Since in driver analysis data are often in the form of time-series, several scholars have used this algorithm in their analysis. We present the details of Dynamic Time Warping in Section 4.2. The technique's goal is to align two sequences of feature vectors by warping the time axis iteratively until an optimal match between the two sequences is found.

In a study conducted by Johnson and Trivedi, MIROAD was designed to detect driving events and driving style. Potentially aggressive driving behaviour is detected and recognized using Dynamic Time Warping (DTW) and data from smartphones, such as accelerometer, gyroscope, magnetometer, GPS, and video (Johnson and Trivedi 2011). Since the length of driving events is not the same, the DTW is a suitable algorithm as it was designed to find the optimal alignment between two signals. In order to detect

whether a driving event was aggressive or not, the DTW algorithm was used to find the closest match between the event signal and different pre-recorded template signals.

In a similar research by Eren et al., a driver's behaviour was detected as safe or unsafe using the DTW algorithm and a Bayesian classifier (Eren et al. 2012). In this study, data such as that from an accelerometer, gyroscope and magnetometer from a typical smart phone was used, and the speed, position angle and deflection from the regular trajectory was computed from the data. Figure 2-44 shows the process of classifying drivers as safe or risky based on data from smartphones using the DTW and Bayesian classifier.

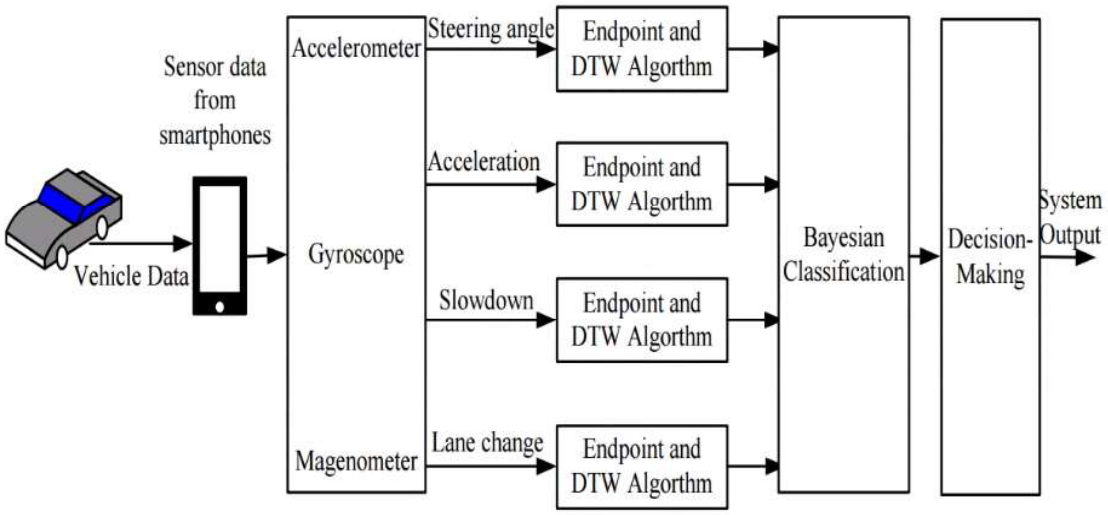

Figure 2-4. Classifying drivers' behaviour into risky or safe based on sensory data from smartphone using DTW and Bayesian classifier.

Various studies have been conducted to explore the individual driving behaviours and cluster them into different clusters. In some of this research, the data used was gathered from simulators. Other research explored data collected from pre-identified driving scenarios, such as aggressive/non aggressive driving, distracted/undistracted driving, etc.

Our work differs in that we look at a specific small portion of driving data (data prior to turning maneuvers) collected from actual drives in an urban area. We are interested in identifying differences among drivers, i.e., clusters, within this short time period. We explore two different approaches. In the first approach, statistical features derived from individual driving signals are used, while in the second, the Dynamic Time Warping

technique is applied to these signals to identify the distance between them. Based on the results of each approach we apply hierarchical clustering analysis to cluster drivers.

Chapter 3

3 Experimental Data

In order to have a better understanding of normal driving behaviour, we need detailed observations on the driver, vehicle and traffic environment. In 2012, Beauchemin, et. al. (Beauchemin et al. 2012) equipped a modern vehicle, with OBD[1] II CAN-bus channels with video cameras, GPS system, cameras to record the driver's head pose and gaze direction. This vehicle, dubbed RoadLAB, was used to collect the driving data from the vehicle's internal network, the environment and the driver. As Figure 3-1 shows, the Driver-Environment-Vehicle (DEV), RoadLAB collected data from a frontal stereo-vision system, faceLAB eye tracker [2], and the CAN-bus interface. This instrumented vehicle can monitor and record the following:

- *Environment:* Front view of traffic environment with two calibrated stereo cameras;
- *Vehicle:* The internal vehicle functions via CAN-bus interface;
- *Driver:* The driver cephalo-ocular behaviour head rotation and gaze direction.

[1] On-board diagnostics (OBD) refers to a vehicle's self-diagnostic and reporting capability

[2] FaceLAB™ 5 (http://www.ekstremmakina.com/EKSTREM/product/facelab/index.html)

Figure 3-1. Equipped car used in RoadLAB to provide Driver-Environment-Vehicle data stream.

The dataset was recorded with 16 different test drivers, consisting 7 males and 9 females, with ages ranging between 20 and 47. The drivers drove normally through a pre-determined path in an urban area inside the city of London, Ontario. Figure 3-2 shows the driving path on Google map. Each driver drove around 28.5 kilometers for about 60 minutes over this route. In total, 3TB of vehicular data was collected over more than 450 kilometers of driving. Table 3-1 shows the general information about each driving experiment conducted.

Figure 3-2. Driving path in urban area- London, Ontario.

Table 3-1. General information about each driving experiment.

Experiments	Driver No.	Age	Gender	Start time	Weather	Temp
1	Driver 1	37	Male	13:15	Sunny	29 ºC
2	Driver 2	37	Male	15:30	Sunny	31 ºC
3	Driver 3	41	Female	12:15	Sunny	23 ºC
4	Driver 4	41	Male	11:00	Sunny	24 ºC
5	Driver 5	37	Female	12:05	Partly cloudy	27 ºC
6	Driver 6	22	Female	13:00	Partly cloudy	21 ºC
7	Driver 7	31	Female	11:30	Sunny	21 ºC
8	Driver 8	21	Male	14:45	Sunny	27 ºC
9	Driver 9	21	Female	13:00	Partly cloudy	24 ºC
10	Driver 10	20	Male	09:30	Sunny	8 ºC
11	Driver 11	22	Female	14:45	Sunny	12 ºC
12	Driver 12	24	Female	11:45	Partly cloudy	18 ºC
13	Driver 13	23	Male	14:45	Partly cloudy	19 ºC
14	Driver 14	47	Female	11:00	Sunny	7 ºC
15	Driver 15	44	Female	14:00	Partly cloudy	13 ºC
16	Driver 16	25	Male	10:00	Partly cloudy	14 ºC

3.1 Data Collection

In this study, we use the data from CAN-bus (Controller Area Network) interface of the vehicle. From the various CAN-bus vehicular data that we have, we extract the most comprehensive and general ones about the vehicle, such as speed, gas pedal pressure, brake pedal pressure and steering wheel angle. These signals were originally sampled at 30 Hz in this experiment. We calculate acceleration based on the speed value, such that in each 1second intervals (30 frames), we calculate V_1-V_0 as the acceleration.

Among all 16 driving experiments, we selected 12 subjects (1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 15, 16) to analyze in this study. The other 4 subjects were not used due to existence of

noise in collected data in some cases and\or because of minor changes in driving path in others.

Figure 3-3 shows examples of 4 driving signals collected during normal urban driving for a period of 300 seconds. From the top, the figures correspond to the speed of the vehicle, gas pedal pressure, brake pedal pressure and steering wheel angle, respectively. As we can see in Figure 3-3, the steering wheel angle value is almost zero in most of the frames although the experiment takes place in urban area. Also, as we might expect, the gas pedal pressure and brake pedal pressure are complementary to each other. In addition, the direct correlation between gas pedal pressure and vehicle speed is obvious.

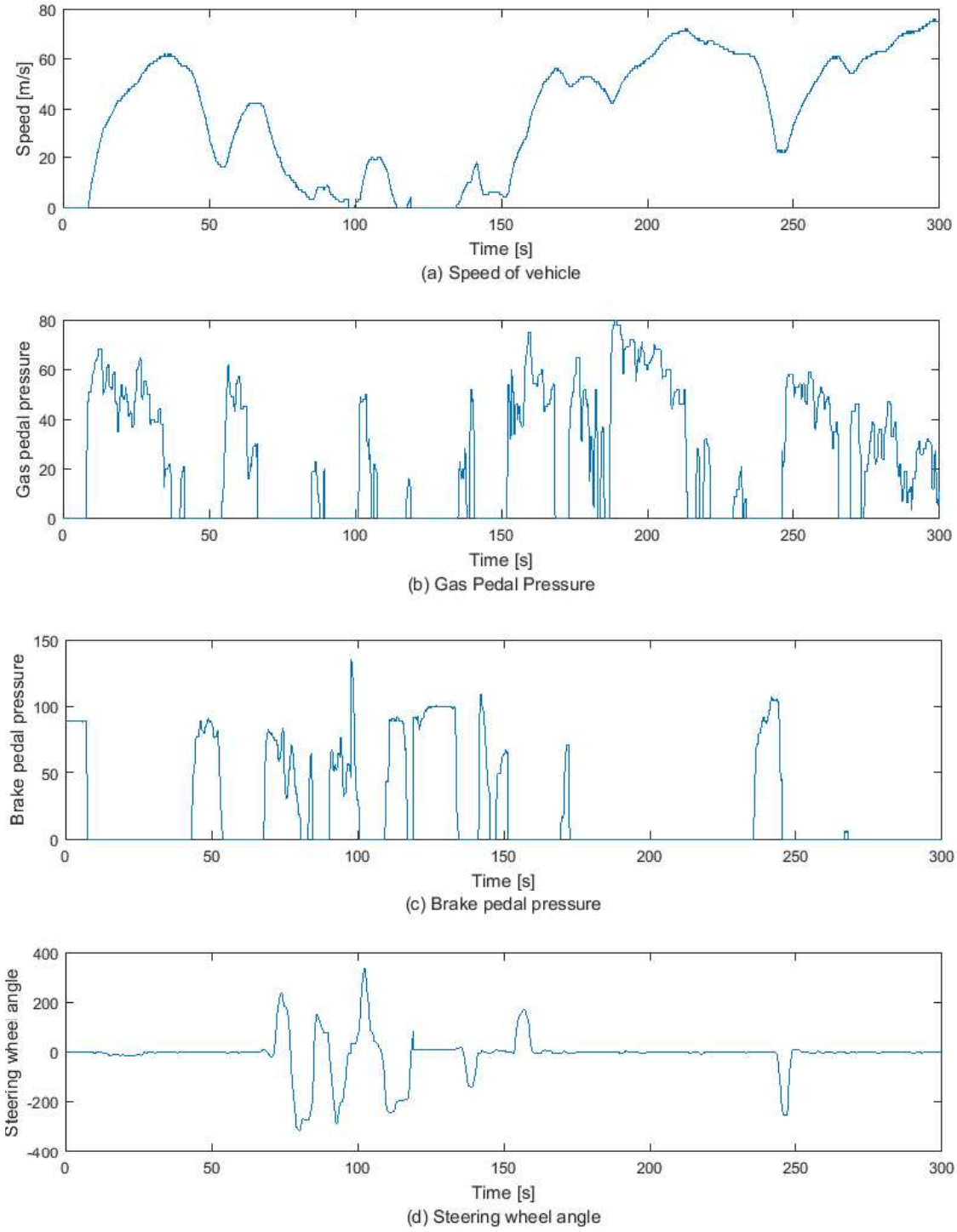

Figure 3-3. Examples of driving behaviour signals collected in 300 seconds of normal driving. (a) speed of vehicle; (b) gas pedal pressure; (c) brake pedal pressure; (d) steering wheel angle.

3.2 Preprocessing

Since the main focus of this study is to categorize drivers based on their driving behaviour considering the pre-turn driving behaviour, we need to extract each of these signals before each turn. In total, we have data for 10 right turns and 7 left turns for 12 drivers consisting of 6 male and 6 female drivers.

As we are analyzing the driver behaviour while driving, we can ignore the time when the vehicle's speed is actually zero (the vehicle stopped). The "stopped state" usually happens right before turning maneuvers, especially before left turns. Heavy traffic, stop signs, traffic light, yield rights, etc. are some common reasons for stopping before turning maneuvers.

Among the not-stopped frames before each turn, we collect three different sets with 150 frames, 300 frames, and 450 frames length (5 seconds, 10 seconds and 15 seconds before each turn respectively).

In the next Chapter the proposed method for clustering drivers based on their driving behaviour before turning maneuvers is discussed and its implementation presented.

Chapter 4

4 Analysis Methods

The main purpose of this study is to explore possible groups of drivers that have similar driving behaviour patterns by clustering individual driver's behaviour based on in-vehicle CAN-Bus signals. These signals include speed, gas pedal pressure, brake pedal pressure, steering wheel angle, and acceleration, collected 5, 10, and 15 seconds before each turning maneuver for different drivers while performing normal driving.

This Chapter highlights the main methods that we applied on these extracted signals in order to come up with the best clustering of driver behaviour. In the first method, different statistical features are extracted from the in-vehicle CAN-Bus signals in order to lower the dimensionality of the data. These statistical features preserve the main characteristics of the corresponding signals and are described in detail in Section 4.1. In Section 4.2, the Dynamic Time Warping (DTW) algorithm (Berndt and Clifford 1994) will be introduced as a distance measure, which we used in our second approach in this study. In this approach, the distance between different time-series signals is calculated using the Dynamic Time Warping algorithm and using these distances, the clustering analysis is performed. These two methods have been used for clustering time series sequences before in other studies, therefore we apply them on the time series driving signals that we have. Time series data are usually high-dimensional data and a specialized distance function is needed to compare them for similarity. Moreover, there might be outliers in these data. Most of machine learning methods, such as K-Means, are designed for low-dimensional spaces with a (meaningful) Euclidean distance. K-Means is not very robust towards outliers, as it puts squared weight on them. Finally, in Section 4.3, hierarchical clustering will be discussed in detail, which is our main clustering method used in this thesis for both approaches.

All the implementation is performed in MATLAB Release 2016a, using the Statistics and Machine Learning Toolbox, and for Dynamic Time Warping a package from MATLAB

Central's File Exchange is used (Wang 2015). Sample implementation code can be found in Appendix C.

4.1 Feature Extraction

The data that we have is a collection of in-vehicle CAN-Bus signals which are time series data. Since the time series data have a unique data structure, it is not easy to directly apply existing data mining tools. In clustering, each time point is often considered a variable and each time series is considered an observation. In time-series data, as the time increases, the number of variables also increases. Therefore, in order to perform clustering we need some feature extraction techniques to summarize the main features of time series data in a significantly lower dimension.

4.1.1 Statistical Feature Extraction

In order to represent a time-series sequence in a lower dimension, we need to transform patterns into features that are considered as a compressed representation. These features construct a high-level representation of the original time-series data.

In this study, for each time series sequence $S_i = \{x_1, x_2, \ldots, x_n\}$, numerous statistical features were calculated to measure different properties of that sequence. Below are details about each of the statistical features extracted and used in this study.

- *Mean*

  The mean μ is the average of the values $\{x_1, x_2, \ldots, x_n\}$ located within a time window of the time-series sequence. It was calculated by:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (4\text{-}1)$$

- *Standard Deviation*

In order to measure how the values $\{x_1, x_2,\ldots, x_n\}$ are spread out within a specific time window of a time-series sequence, the standard deviation σ is calculated as:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \quad (4\text{-}2)$$

- *Kurtosis*

Kurtosis *Ku* is a measure of the "peakedness" of the probability distribution of the values $\{x_1, x_2,\ldots, x_n\}$ within a specific time window of a time-series sequence. It was calculated as:

$$Ku = \frac{\mu_4}{\sigma^4} \quad (4\text{-}3)$$

where $\mu_4$is the 4th moment about the *mean*, and is given as:

$$\mu_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^4 \quad (4\text{-}4)$$

- *Skewness*

In order to measure the asymmetry of the sequence of data $\{x_1, x_2,\ldots, x_n\}$, the skewness was calculated as:

$$Sk = \frac{\mu_3}{\sigma^3} \quad (4\text{-}5)$$

where μ_3 is the 3rd moment about the *mean*, and calculated as:

$$\mu_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^3 \quad (4\text{-}6)$$

The above mentioned basic statistical functions were calculated for each of the 5 different signals (speed, gas pedal pressure, brake pedal pressure, steering wheel angle and acceleration) before each turn, which results in 4x5=20 statistical features representing the characteristics of signals corresponding to a single driver's behaviour before each turn.

Having these feature vectors for each driver before each turning maneuver, we could now compute the distances between drivers and attempt to cluster them based on these statistical features representing their driving behaviour prior to turns.

4.2 Dynamic Time Warping

As discussed in previous section, the general characteristics and dimensionality of the time series data are different. Temporal sequences have different characteristics considering normal feature based data, which make the process of comparing sequences more challenging. One algorithm that is widely use in order to find similarities between time series sequences is Dynamic Time Warping (Berndt and Clifford 1994).

Dynamic time warping (DTW) is a time series alignment algorithm developed originally for speech recognition (Sakoe and Chiba 1978). In 1994 Berndt and Clifford introduced the technique, dynamic time warping (DTW), to the database community (Berndt and Clifford 1994). The technique's goal is to align two sequences of feature vectors by warping the time axis iteratively until an optimal match between the two sequences is found. Figure 4-1 shows an optimal alignment between two time dependent sequences.

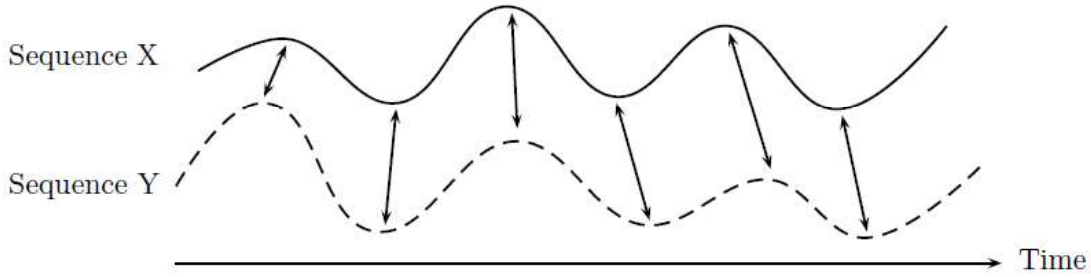

Figure 4-1. Alignment of two time-series sequences. Aligned points are indicated by arrows (Meinard Müller 2007).

The objective of DTW is to compare two (time-series) sequences $X = (x_1, x_2, \ldots, x_N)$ of length $N \in \mathbb{N}$ and $Y = (y_1, y_2, \ldots, y_M)$ of length $M \in \mathbb{N}$. To be able to compare two different instances x, y, we need a *local cost measure*, sometimes also referred to as *local distance measure*. Typically, $c(x, y)$ is small (low cost) if x and y are similar to each other, and otherwise $c(x, y)$ is large (high cost). Evaluating the local cost measure for each pair of elements of the sequences X and Y, one obtains the *cost matrix* (or *distance matrix*) $C \in \mathbb{R}^{N \times M}$ defined by $C(n,m) := c(x_n, y_m)$. Then the goal is to find an alignment between X and Y having minimal overall cost (Meinard Müller 2007).

4.2.1 Constraints for the optimal path

In order to reach this goal and find the best match between two sequences, we need to find a path through the cost matrix which minimizes the total cost between them. The overall cost is the minimum of all possible path between troughs in the cost (distance) matrix. Obviously, the number of possible paths through the matrix between two considerably long sequences can be enormous. Exploring every possible path would lead to a computational complexity that is exponential in the lengths of N and M. In order to lower the order of computational complexity to $O(NM)$, several optimisations and constraints are applied when using DTW:

- *Monotonicity condition:* With this condition, during the construction of a path, indexes i and j never decrease, they either stay the same or increase.

- *Continuity condition:* This condition limits *i* and *j* indexes to increase by at most 1 in each step of the path.
- *Boundary condition:* The path starts at $i = j = 0$ and ends at $i = M$ and $j = N$, where M and N are the number of instances in the first and second sequences respectively.
- *Warping window condition:* It is unlikely for the path to wander very far from the diagonal. The window width determines the distance that the path is allowed to wander.

- *Slope constraint condition:* This condition would prevent subsequences with very different lengths to match by limiting the number of steps allowed in the same direction (horizontal or vertical). The condition is expressed as a ratio p/q, where p is the number of steps allowed in the same (horizontal or vertical) direction. After p steps in the same direction, the algorithm is not allowed to step further in the same direction before stepping at least q time in the diagonal direction.

Figure 4-2 shows the smoothed speed signal of a vehicle 10 seconds before a specific right turn when driven by two different drivers (Driver8 and Driver12). The corresponding optimal warping path p* between two signals is demonstrated in Figure 4-3 (the white line). The gray area in this figure shows the Cost Matrix between two sequence within the warping window around the diagonal and the white line indicate the optimal warping path. Figure 4-4 compares the original signals (a), with warped ones (b) which are perfectly aligned.

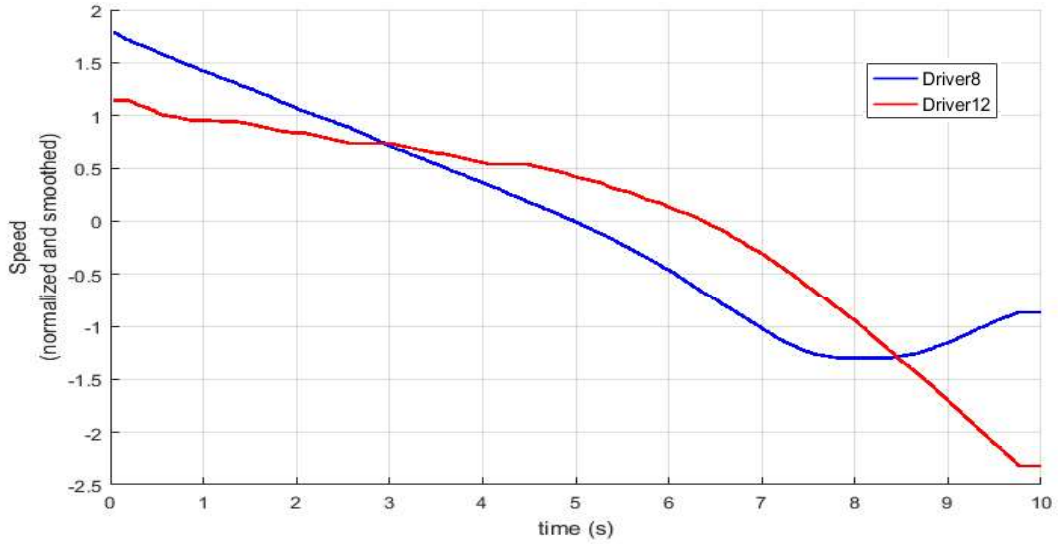

Figure 4-2. Speed signal of two drivers' behaviour 10 seconds before a specific turn.

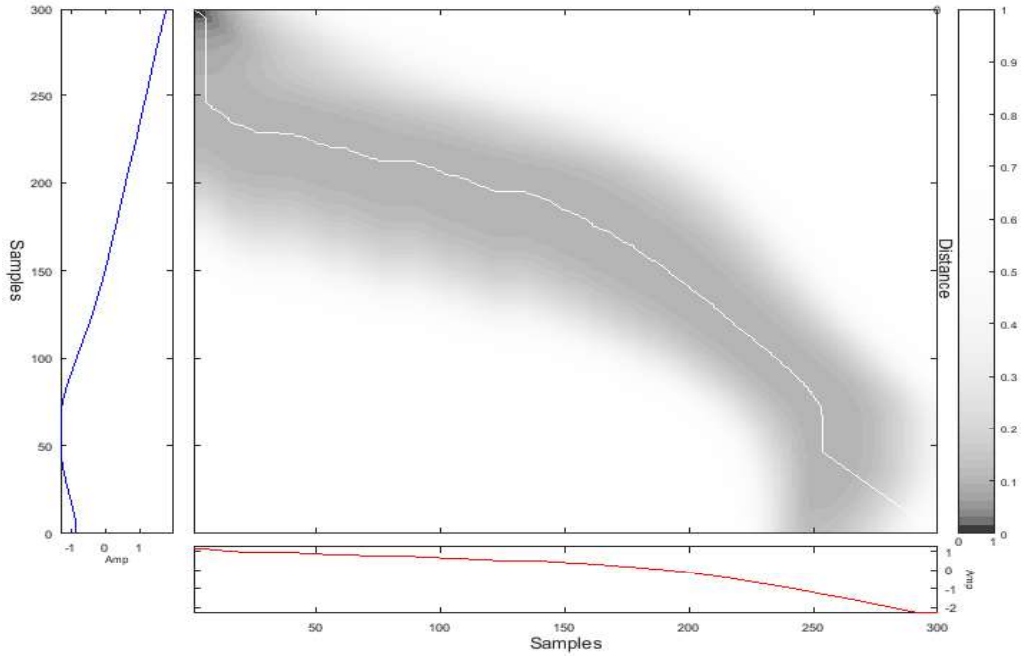

Figure 4-3. Accumulated cost matrix *D* with optimal warping path $p*$ (*white line*).

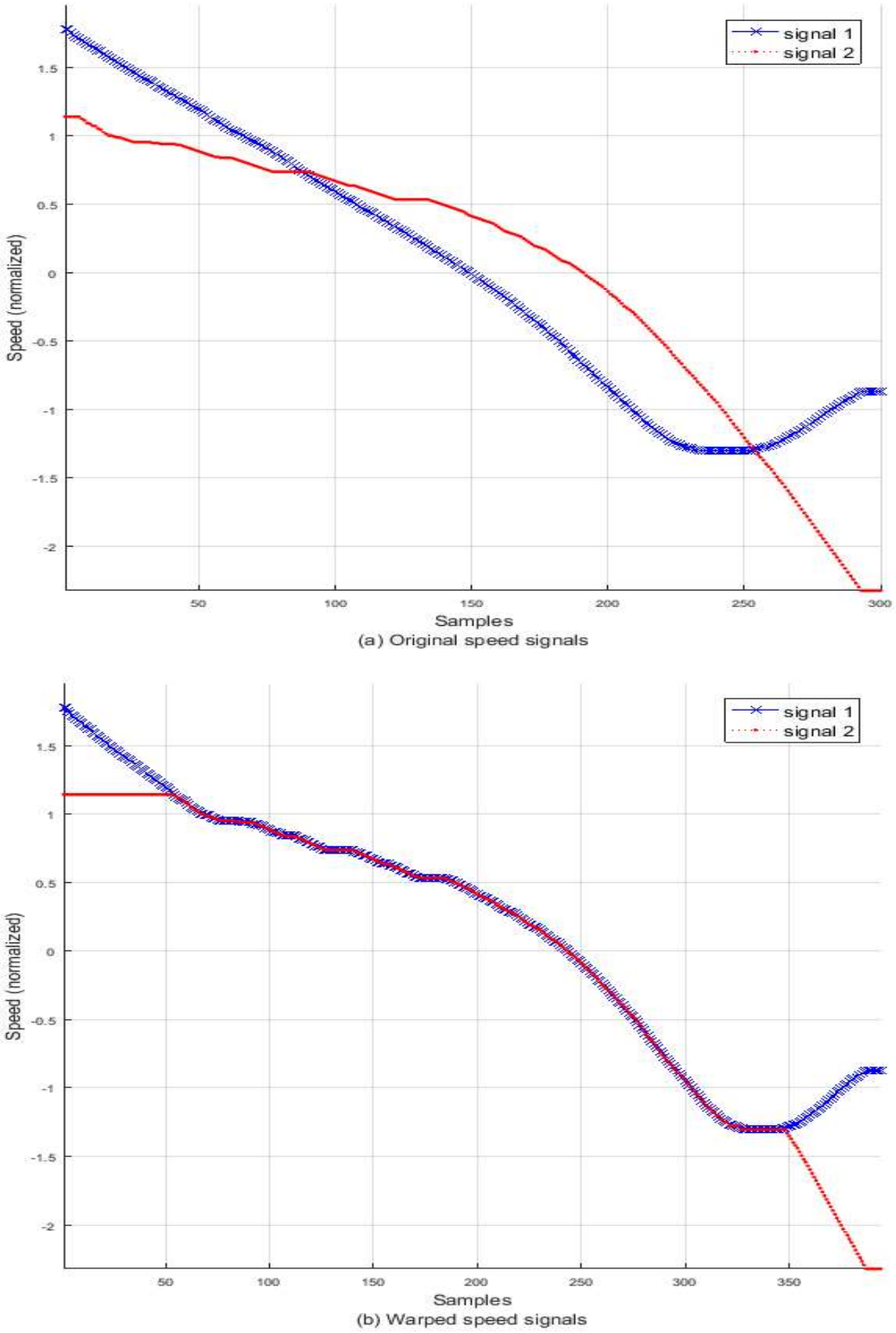

Figure 4-4. (a) Original speed signal (b) Warped speed signal.

In this study, we apply the Dynamic Time Warping package, implemented by (Wang 2015) on different CAN-bus signals, such as speed, gas pedal pressure, brake pedal pressure, steering wheel angle, and acceleration, for 5, 10, and 15 second of driving (excluding zero-speed frames) before each turn. In order to analyze the turning behaviour better, we first analyse all the turns and then left turns and right turns are analyzed separately. As a result, for each category of all turns, right turns, or left turns, we come up with three 12x12 distance matrices which shows the distance between each of the 12 drivers for driving behaviour of 5, 10, and 15 seconds before turns, respectively. Therefore, overall we have nine distance matrixes which are our references for the distance measure used in hierarchical clustering (three distance matrixes for all turns, three for left turns, and three for right turns). Using these distance matrices, we can now apply hierarchical clustering analysis to categorize the drivers based on their driving signals similarities. In the next section, hierarchical clustering and various distance metrics are discussed in detail.

4.3 Hierarchical Clustering

One of the most famous clustering methods is hierarchical clustering, which has been applied in many domains. Hierarchical clustering (also known as hierarchical cluster analysis or HCA) is a method of cluster analysis which attempts to build a hierarchy of clusters. In the process of clustering, a dendogram[1] is generated to visualize the result of clustering, representing the nested clustering of patterns and similarity levels at which groupings change.

Two different strategies are applied in hierarchical clustering:

[1] A dendogram is a tree diagram which illustrate the arrangement of the clusters produced by hierarchical clustering.

- *Agglomerative:* Also known as a "Bottom-up" approach. This approach starts with putting each observation in its own cluster and merge the clusters as we go up the hierarchy

- *Divisive:* Also known as "Top-down" approach. This approach starts with one cluster containing all the observation splits down into different clusters as we go down the hierarchy

Figure 4-5 shows these two general approaches to hierarchical clustering. In this study we apply agglomerative clustering approach.

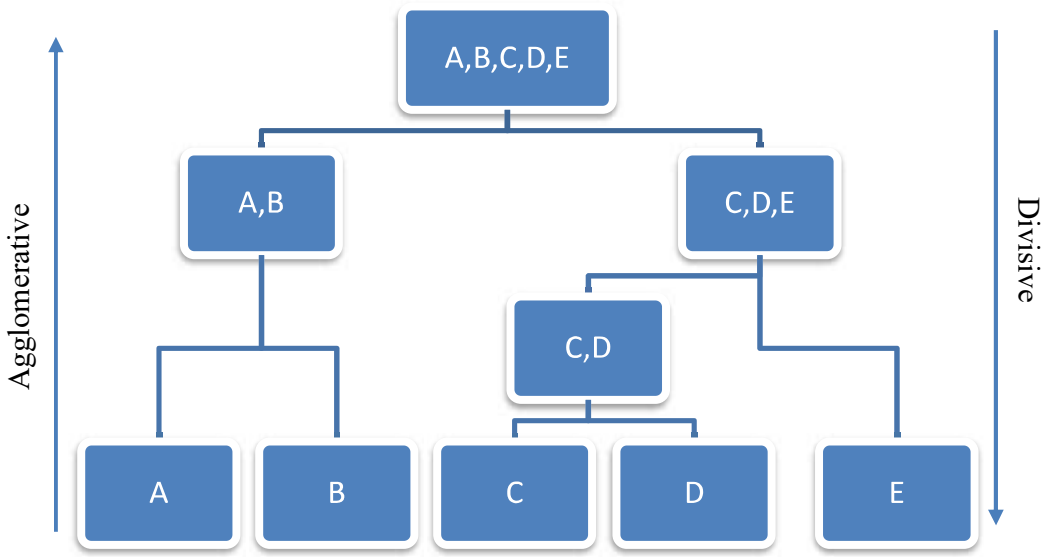

Figure 4-5. Agglomerative versus Divisive approach in hierarchical clustering.

4.3.1 Cluster dissimilarity

In order to see which two clusters should be merged in an agglomerative approach or when a cluster should be split in divisive approach, a distance measure (metric) between observations and a linkage criterion is required.

4.3.1.1 Distance Metric

It is really important to choose an appropriate metric as the whole shape of clusters relate directly to the distance metric. This is because two instances may be close based on one distance metric and farther away according to another.

In the following, some common distance metrics for hierarchical clustering are presented. Consider p = (p_1, p_2,..., p_n) and q = (q_1, q_2,..., q_n), and the distance between p and q , $d(p,q)$, can be calculated by the following distance metrics:

- *Euclidean distance*

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \quad (4\text{-}7)$$

- *Squared Euclidean distance*

$$d(p,q) = \sum_{i=1}^{n}(p_i - q_i)^2 \quad (4\text{-}8)$$

- *Manhattan distance*

$$d(p,q) = \sum_{i=1}^{n}|p_i - q_i| \quad (4\text{-}9)$$

- *Maximum distance*

$$d(p,q) = \max_i |p_i - q_i| \quad (4\text{-}10)$$

- *Mahalanobis distance*

$$d(p,q) = \sqrt{(p-q)^T S^{-1}(p-q)} \quad (4\text{-}11)$$

where S is the covariance matrix. The covariance matrix is a matrix whose element in the i, j position is the covariance between the i^{th} element of vector p and j^{th} element of vector q.

- *Other metrics*

  There are many distance metrics that can be used. In this study, Euclidian distance is used to find the distance between statistical features of observation signals in our first approach and in the second approach the output of Dynamic Time Warping is used as the distance matrix of drivers directly.

The Mahalanobis Distance is one metric that has been frequently used in cluster analyses. However, in order to use the Mahalanobis Distance, the number of observations needs to be more than the number of features. Since we have just 12 drivers in this study, we could not use this metric. Among the other metrics, the Euclidean Distance seems to be more useful for clustering analysis in Statistical Feature Extraction approach.

4.3.1.2 Linkage Criterion

In order to merge or split two clusters we need to be able to measure the distance of two sets of observations. A linkage criterion measures this distance as a function of distance metrics *d*. Here *D(P,Q)* is the merging cost of combining clusters *P* and *Q*. In the following, some commonly used linkage criteria between two sets of observations P and Q are introduced:

- *Complete-linkage*

$$D(P,Q) = \max\{d(p,q) : p \in P, q \in Q\} \quad (4\text{-}12)$$

- *Single-linkage*

$$D(P,Q) = \min\{d(p,q) : p \in P, q \in Q\} \quad (4\text{-}13)$$

- *Average linkage*

$$D(P,Q) = \frac{1}{|P||Q|}\sum_{p\in P}\sum_{q\in Q} d(p,q) \quad (4\text{-}14)$$

- *Centroid linkage*

$$D(P,Q) = ||c_P - c_Q|| \quad (4\text{-}15)$$

where c_P and c_Q are the centroids of clusters P and Q respectively.

- *Ward linkage*

This method minimizes the total within-cluster variance.

$$D(P,Q) = \sum_{i\in P\cup Q} \| \vec{x}_i - \vec{m}_{P\cup Q} \|^2 - \sum_{i\in P} \| \vec{x}_i - \vec{m}_P \|^2 - \sum_{i\in Q} \| \vec{x}_i - \vec{m}_Q \|^2 = \frac{n_P n_Q}{n_P + n_Q} \| \vec{m}_P - \vec{m}_Q \|^2 \quad (4\text{-}16)$$

where $\vec{m}_P$ is the center of cluster P, and n_P is the number of instances in it.

Since, in this analysis, our goal is to find clusters with minimum variation, the “ward” method as linkage criteria seems to fit our needs.

As discussed in the previous sections of this Chapter, in this research two different approaches are used to measure the distances between in-vehicle CAN-bus signals of various drivers’ driving behaviour before turning events. Hierarchical clustering can have two different set of inputs. The input can be the observations, distance metric and linkage criterion (as we need in our first approach with statistical feature vectors), or it be the distance matrix between observations and the linkage criterion (which we need in our second approach with the DTW technique).

Based on the extracted statistical features discussed in Section 4.1, hierarchical clustering using the Euclidean distance as a distance metric and the Ward method as linkage criteria was used. This results in 9 cluster trees clustering 12 drivers based on 5, 10, and 15 seconds before all turns, before right turns and before left turns separately. In addition, as discussed in Section 4.2, DTW was also used as a distance metric and results in a distance matrix that can be used to cluster the drivers based on similarities between in-vehicle signals before turning events. In this approach, we used the DTW as a distance metric and the ward linkage method as the linkage criteria for the hierarchical clustering. This approach, as well, results in 9 cluster trees, three of them correspond to 5, 10, and 15 seconds before all turns, three for before right turns, the other three for before left turns.

The results and corresponding clustering figures for both approaches are shown in the next Chapter.

Chapter 5

5 Analysis of Results and Discussion

5.1 Introduction

To evaluate the proposed method described in the previous chapter, an experiment was conducted to cluster typical behaviour of drivers considering how they drive before turning events. As mentioned in Chapter 3, in this study our data consists of 10 different right turns and 7 different left turns for 12 drivers, among them there are 6 male and 6 female drivers.

We collect three sets of data with different numbers of driving frames in each set, consisting of 150, 300, and 450 frames before each turn (equal to 5, 10, and 15 seconds). Choosing the correct number of frames is an important factor in analyzing the result. A small number of frames may not reflect signal changes needed for a proper clustering, on the other hand, since in our experiment some of turning maneuvers are close to each other, choosing a large number of frames may contain frames from the previous turn. A set of 300 frames seems to be a reasonable, although we perform our analysis on 150 frames and 450 frames too.

Five different CAN-bus signals consisting of speed, gas pedal pressure, brake pedal pressure, steering wheel angle, and acceleration are extracted from each set of frames for each driver before turning events. Two different approaches were used to analyze this data in order to explore clustering drivers into different groups; clustering using statistical features and clustering using Dynamic Time Warping. In the next section we provide an overview of both approaches.

5.2 General Overview

The general overview of the two approaches that we apply to our data is illustrated in Figure 5-1. The approach using Statistical Feature Extraction is illustrated in Figure 5-1.a. and the Dynamic Time Warping approach is shown in Figure 5-1.b.

5.2.1 Statistical Feature Approach

As is shown in Figure 5-1.a, we first extract the 5 CAN-Bus time-series signals over 17 turns for each of the 12 drivers. From each time-series signal that we have, we extract four statistical features (*Mean, STD, Kurtosis, Skewness*), representing driving behaviour of each driver before each turn. Then in step a.1, we calculate the average of these features over all turns. The result is a 5x4 matrix for each driver. Then in step a.2., for each driver we convert its 5x4 matrix to a feature vector of size 20. Doing that for all 12 drivers we end up with 12x20 feature matrix. Then we apply HCA on the result of all drivers' feature vectors using the "ward" linkage criteria and "Euclidean distance" as the distance metric. The result of the HCA is a dendogram clustering drivers based on statistical features extracted from pre-turns driving behaviour. We do this for each of the 150, 300 and 450 frame sequences.

5.2.2 DTW Approach

The Dynamic Time Warping approach is illustrated in Figure 5-1.b. We have the same data as we had in statistical approach, containing 12 drivers and for each driver we have 5 CAN-Bus time-series signals over 17 turns. As mentioned in step b.2. we calculate the average distance of all corresponding signals between each pair of drivers over all turns using DTW. This will result a symmetric 12x12 distance matrix which indicates the distance between each pair of drivers. In the next step, we apply HCA on this matrix using the "ward" linkage criteria to achieve the clustering dendogram.

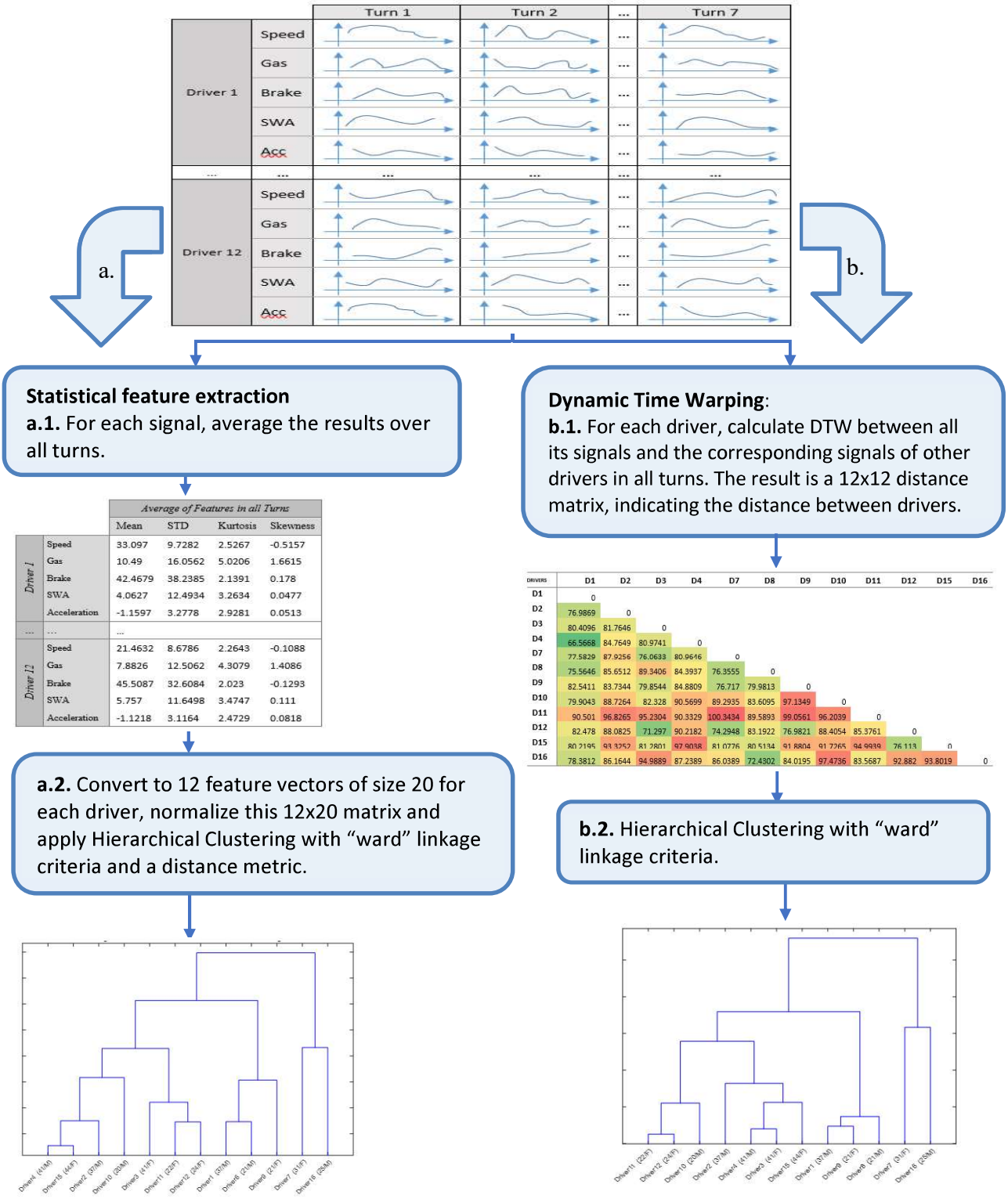

Figure 5-1. An overview of the two approaches. a. The Statistical feature extraction approach. b. The DTW approach.

5.3 Analysis of Statistical Features

As mentioned in Chapter 4, in order to lower the dimensionality of our data we extract 4 statistical features containing *Mean*, *STD*, *Kurtosis*, and *Skewness* from each of 5 CAN-bus signals and use them as a descriptor for the corresponding signal. Using these extracted features, we can find the distance between each pair of signals for different drivers. Combining all the statistical features from all the signals for each driver results in a 2-dimentional matrix of size 5 x 4 that corresponds to the driving behaviour before each turning event. Since we have 10 right turns and 7 left turns in our experiment, we obtain three 3-dimentional matrixes representing each driver. One 3D matrix corresponds to right turns with size 5 x 4 x 10, one for left turns with size 5 x 4 x 7, and one for all turns of size 5 x 4 x 17. Figure 5-2 shows the general structure of a 5 x 4 x 7 3D matrix indicating feature values of signals corresponds to all left turns of a specific driver.

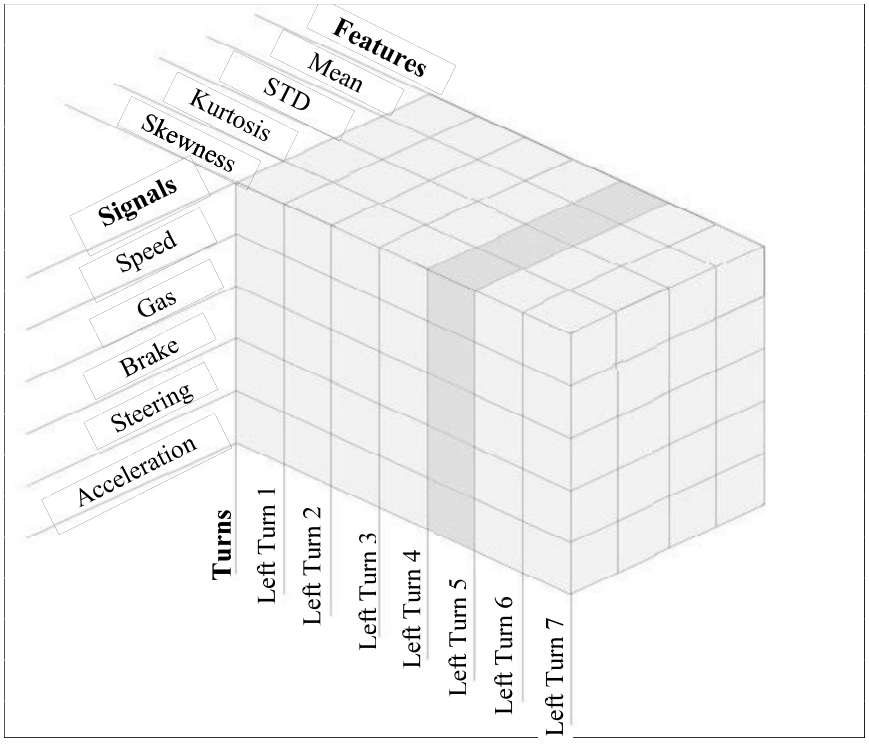

Figure 5-2. The general structure of 3-dimentional 5x4x7 matrix of the 4 extracted statistical features from 5 CAN-Bus signals in 7 left turns of a specific driver behaviour. The highlighted 2D matrix corresponds to statistical features of fifth left turn as an example.

We now calculate the average value of each statistical features for each signal over all the turns. So we would end up with a 5x4 feature matrix for each driver. Then we convert this matrix to a feature vector of size 20 for each driver, representing the behaviour of the driver over all the turns. Combining all the 12 drivers feature vector would result a 12x20 matrix of drivers-features.

Because of the variation in the range of each of these statistical features, we normalize each column of this matrix using a feature scaling method in order to come up with more reasonable distances. The feature scaling method used to bring all features into the range [0,1] was performed as:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (5\text{-}1)$$

where X_{max} is maximum value and X_{min} is the minimum value in the set.

In our case, we consider all the values of a specific feature (e.g. Maximum speed) from all 12 drivers, and perform feature scaling on them. Feature scaling not only keeps the differences between values of same features but also scales all the features in the same range to prevent domination of specific features with large variations.

Table 5-1 shows an example of real values of average statistical features extracted from each of the CAN-Bus signals of driver1's driving behaviour 450 frames before all turns, and Table 5-2 shows the normalized values in the same situation. Note that feature scaling of a feature is done considering the values of that feature from all the drivers. More actual values of average statistical features of each driving signal can be found in Appendix A.

Table 5-1. An example of statistical features values extracted from signals of driver1's driving behaviour, 450 frames before all turns.

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 1	Speed	33.097	9.7282	2.5267	-0.5157
	Gas	10.49	16.0562	5.0206	1.6615
	Brake	42.4679	38.2385	2.1391	0.178
	SWA	4.0627	12.4934	3.2634	0.0477
	Acceleration	-1.1597	3.2778	2.9281	0.0513

Table 5-2. An example of normalized* statistical features values extracted from signals of driver1's driving behaviour, 450 frames before all turns.

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 1	Speed	0.8950	0.4421	0.0169	0.1183
	Gas	0.6253	0.9190	0.3880	0.8027
	Brake	0.8530	1.000	0.1163	0.3925
	SWA	0.5974	0.3461	0.0063	0.1484
	Acceleration	0.3617	0.9999	0.9999	0.9488

** Normalization of each feature is done considering corresponding feature from all drivers*

After doing feature scaling on the extracted features, the 'ward' method as linkage criteria and 'Euclidean distance' as the distance metric is used for hierarchical cluster analysis. In order to find any possible considerable differences in driving behaviour prior to right turns or left turns, we not only analyze the all pre-turns but also analyze pre-left-turns and pre-right-turns driving behaviour separately. The results are shown in the next three sub-sections.

5.3.1 All turns

We perform Hierarchical Clustering Analysis (HCA) on 150, 300, and 450 frames before all turns for 12 drivers using their statistical features. The dendogram related to HCA on 150 frames before all turns is shown in Figure 5-3 as an illustration. Other dendograms related to 300 and 450 frames before turns can be found in Appendix B.

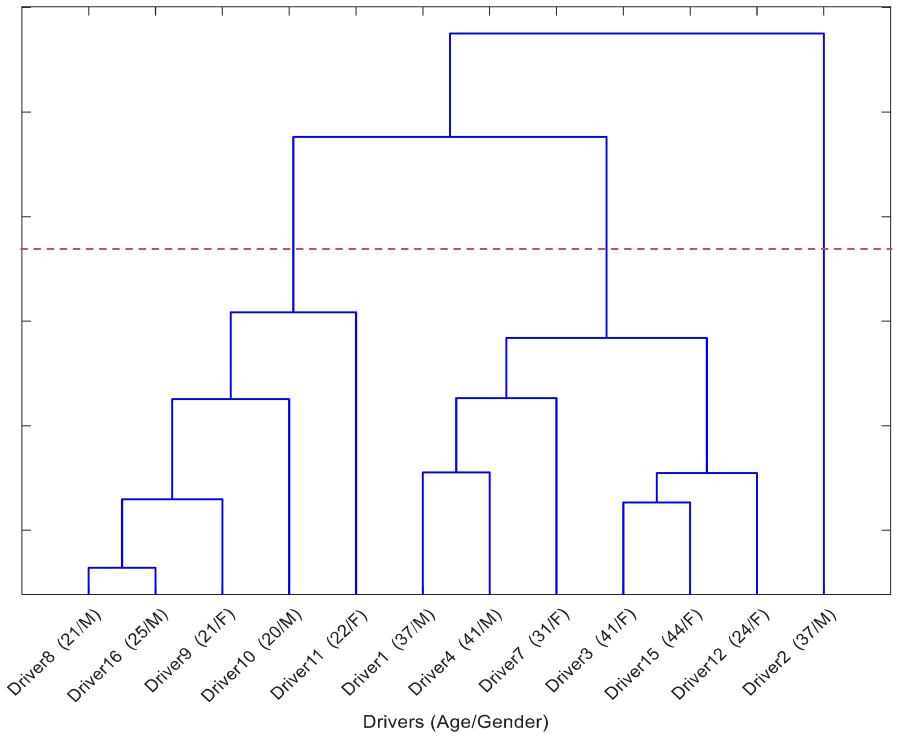

Figure 5-3. HCA based on statistical features considering 150 driving frames before all turns.

For all of our dendograms in analyzing all pre-turns using statistical features, we can identify two clusters of drivers (the red dashed line indicates the threshold). In order to investigate each cluster, we calculate the centroid of each the statistical features in each cluster. In this case, Cluster 1 include drivers “8, 16, 9, 10, 11” and Cluster 2 contains drivers “1, 4, 7, 3, 15, 12”. The centroid of two clusters obtained from HCA on 150 pre-turns frames related to Figure 5-3 is presented in

Table 5-3. The similar centroid values of clusters in other clustering analyses using 300 and 450 frames before all turns can be found in Appendix B.

Table 5-3. Centroid values of statistical features in clusters from HCA on 150 driving frames before all turns.

	Cluster 1 "8, 16, 9, 10, 11"					*Cluster 2 "1, 4, 7, 3, 15, 12"*				
	Speed	Gas	Brake	SWA	Acc	Speed	Gas	Brake	SWA	Acc
Mean	0.405	0.6087	0.3294	0.8471	0.6988	0.8333	0.2997	0.8714	0.6059	0.2686
STD	0.0559	0.566	0.3839	0.0989	0.2721	0.2149	0.3254	0.645	0.1457	0.4638
Kurtosis	0.0497	0.1338	0.3225	0.1786	0.5397	0.0572	0.4829	0.0555	0.1136	0.5264
Skewness	0.26	0.1772	0.5131	0.4522	0.6386	0.2343	0.5872	0.1442	0.555	0.6853

As shown in

Table 5-3, comparing the centroid values of cluster 1 and 2 we can see that they are mostly different. To identify the actual difference between centroids of these two clusters, we convert each clusters' centroids matrix to a vector of size 20 and calculate the Angle and the R-Value[1] between these two vectors. The Angle and R-Value between the two clusters in each experiment using 150, 300, and 400 frames are shown in Table 5-4. This table also shows the driver number and the number of Male and Female drivers in each cluster.

Table 5-4. Angle and R-Value between centroid vectors of two clusters result from 150, 300, and 450 frames before all turns.

	Cluster 1 (Drivers No.)	***Cluster 2 (Drivers No.)***	***Cluster 1 M/F***	***Cluster 2 M/F***	***Angle (degrees)***	***R-Value***
All Turns 150	8,9,**10**,11,16	**1,3,4**,7,12,15	3M, 2F	2M, 4F	35	0.32
All Turns 300	1,7,8,**10**,15,16	**3,4**,**9,11**,12	4M, 2F	1M, 4F	33	0.35
All Turns 450	2,7,**10**,12,15	**1,3,4**,8,**9,11**,16	2M, 3F	4M, 3F	35	0.06

The Angle value indicates the angle between the two vectors, and the R-Value shows the correlation between each pair of vectors. The R-Values and Angles show that the clusters are not particularly correlated. Moreover, there is some consistency in each cluster,

[1] R-value is the correlation coefficient, measures the strength and direction of a linear relationship between two variables. The value of r is always between +1 and –1.

Cluster 1 and Cluster 2, across all three timeframes, that is, there is some consistency in membership in these clusters.

Among all the five signals that we investigate through this research, the raw value of all signals are always positive except for the Steering Wheel Angle (SWA) and Acceleration. Since these signals can have both negative and positive values (indicating steering to right or left, and accelerating or decelerating), it is important to see how zero value in these signals changed after normalizing the signals between 0 and 1. In a normal distribution of these signals we expect zero to be around 0.5, if it is less than 0.5, then we had more positive values in the data set, if it is more than 0.5, the we have more negative values, below 0 means that all values are positive, and more than 1 means all values are negative. The zero value in SWA and Acceleration for each experiment is presented in Table 5-5. This table also include the mean value of each signals in each cluster.

Table 5-5. Mean value of all signals of two clusters results from performing HCA on 150, 300, and 450 frames before all turns.

All Turns			*Cluster 1*					*Cluster 2*				
	SW = 0	*Acc = 0*	*Mean Speed*	*Mean Gas*	*Mean Brake*	*Mean SW*	*Mean Acc.*	*Mean Speed*	*Mean Gas*	*Mean Brake*	*Mean SW*	*Mean Acc.*
150 Frms	0.2440	1.3176	0.4050	0.6087	0.3294	0.8471	0.6988	0.8333	0.2997	0.8714	0.6059	0.2686
300 Frms	0.1928	1.7536	0.5445	0.7271	0.5316	0.5588	0.7914	0.6940	0.3181	0.8608	0.6107	0.3103
450 Frms	0.0888	1.3403	0.6327	0.8942	0.5349	0.5188	0.6326	0.5368	0.3072	0.8215	0.5486	0.2989

Comparing the mean value of each of the signals, we can see that there are obvious differences between Cluster 1 and Cluster 2. Based on Table 5-5 we can see that zero value of SWA after normalization is way less than 0.5, which indicates that prior to all turns, drivers have been steering more to right than left. There are more right turns in the route, hence the mean Steering Wheel values are skewed to right.

Also, the zero value of Acceleration after normalization is more than 1 in all three timeframes, which means that prior to all turns, all drivers decelerate, therefore the higher value in this signal means lower deceleration (smaller negative acceleration). Considering the mean value of the Acceleration signal, it is obvious that drivers in Cluster1 have more

values (lower deceleration) than Cluster2 (which have higher deceleration). Which means that drivers in Cluster2 slow down more quickly than those in Cluster1, and as a result they push the brake harder and they have lower pressure on these gas pedal. This behaviour leads us to classify drivers in Cluster2 as “more aggressive” when compared to the drivers of Cluster1, which we classify as “moderate” drivers.

So to sum up, for 150 and 300 timeframes the two clusters have these behaviours approaching turns:

- *Cluster 1 (Moderate Drivers):* Moderate speed, some gas pressure approaching turn, gentle braking, and gradual deceleration;
- *Cluster 2 (Aggressive Drivers):* Higher speed approaching turns, harder braking and more rapid deceleration.

For 450 timeframes, signal values start to level out, but differences between two clusters are still noticeable.

5.3.2 Right Turns

As with all turns, we perform Hierarchical Clustering Analysis (HCA) on 150, 300, and 450 frames looking at the time before right turns for the 12 drivers. The dendogram from the HCA on 150 frames before right turns is shown in Figure 5-4. Other dendograms related to 300 and 450 frames before right turns can be found in Appendix B.

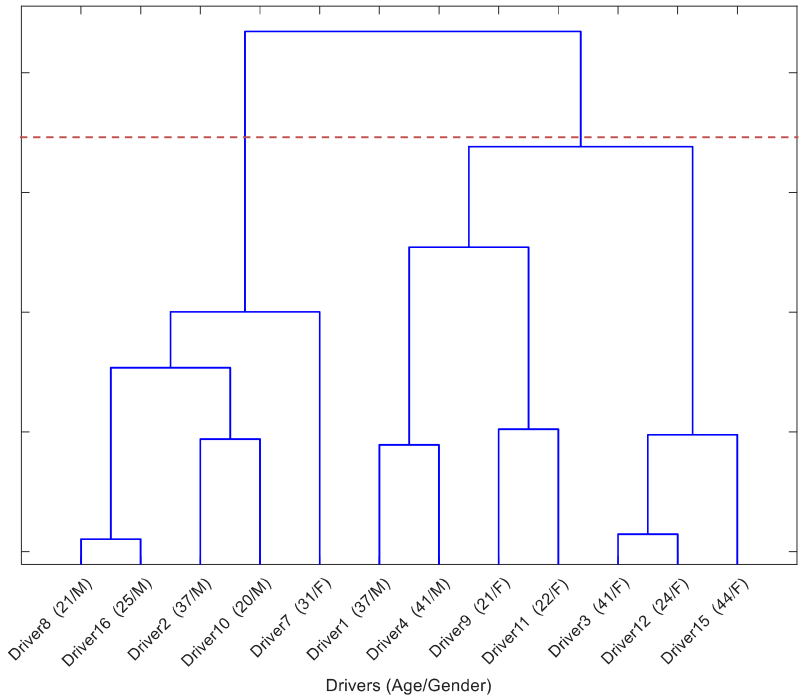

Figure 5-4. HCA based on statistical features considering 150 driving frames before right turns.

For all of our dendograms from analyzing right turns using statistical features, we can identify two clusters of drivers (the red dashed line indicate the threshold). In order to investigate each cluster, we calculate the centroid of each statistical features in each cluster. According to result shown in Figure 5-4, Cluster 1 included drivers "8, 16, 2, 10, 7" and Cluster 2 contains drivers "1, 4, 9, 11, 3, 12, 15". The centroid of two clusters obtaining from HCA on 150 frames before right turns (related to Figure 5-4) is presented in Table 5-6. The similar centroid values of clustering in 300 and 450 timeframes before right turns can be found in Appendix B.

Table 5-6. Centroid values of statistical features in clusters from HCA on 150 driving frames before right turns.

	Cluster 1 "8, 16, 2, 10, 7"					*Cluster 2 "1, 4, 9, 11, 3, 12, 15"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.4361	0.8036	0.2299	0.7078	0.8814	0.7089	0.2972	0.7459	0.7424	0.2996
STD	0.197	0.7463	0.5817	0.46	0.4311	0.616	0.3554	0.5632	0.1549	0.5959
Kurtosis	0.3169	0.1119	0.3667	0.377	0.6081	0.3392	0.2931	0.4624	0.56	0.5932
Skewness	0.655	0.2486	0.7861	0.4852	0.5519	0.4272	0.5687	0.5543	0.5208	0.7837

Based on the centroid values in Table 5-6, we can see that the centroid values of signals related to Cluster1 and Cluster2 are mostly different. As with our analysis of all turns, in order to identify the actual difference between centroids of these two clusters, we convert each clusters' centroids 4x5 matrix to a vector of size 20 and calculate the Angle and the R-Value between these two vectors. The Angle and R-Value between two clusters in each experiment using 150, 300, and 400 frames are shown in Table 5-7.

Table 5-7. Angle and R-Value between centroid vectors of two clusters resulting from 150, 300, and 450 frames before right turns.

Right Turns	*Cluster 1 (Drivers No.)*	*Cluster 2 (Drivers No.)*	*Cluster 1 M/F*	*Cluster 2 M/F*	*Angle (degrees)*	*R-Value*
150 Frames	**2,7,8,10,16**	1,**3,4,9,11,**12,15	4M, 1F	2M, 5F	32	-0.15
300 Frames	1,**2,7,8,10**,12,15,**16**	**3,4,9,11**	5M, 3F	1M, 4F	37	-0.29
450 Frames	**2,7,8,10**,12,15,**16**	1,**3,4,9,11**	4M, 3F	2M, 4F	34	-0.35

Again, we can see some consistency among drivers within each of Cluster 1 and Cluster 2 across the different timeframes. Here, we can see that the membership in the clusters is fairly consistent across the different timeframes, which shows some consistency in the membership in both clusters.

Table 5-8 contains the mean value of signals in different timeframes before right turns. The normalized zero value for Steering Wheel (SWA) and the Acceleration for each analysis is also presented in this table.

Table 5-8. Mean value of all signals of two clusters result from performing HCA on 150, 300, and 450 frames before right turns.

Right Turns			*Cluster 1*					*Cluster 2*				
	SW = 0	*Acc = 0*	*Mean Speed*	*Mean Gas*	*Mean Brake*	*Mean SW*	*Mean Acc.*	*Mean Speed*	*Mean Gas*	*Mean Brake*	*Mean SW*	*Mean Acc.*
150 Frms	0.4654	1.3716	0.4361	0.8036	0.2299	0.7078	0.8814	0.7089	0.2972	0.7459	0.7424	0.2996
300 Frms	0.3700	1.4805	0.5384	0.6990	0.4022	0.5162	0.7393	0.6814	0.2368	0.8092	0.7114	0.1589
450 Frms	0.2565	1.0368	0.3910	0.6713	0.5391	0.4375	0.6037	0.7040	0.4322	0.7309	0.5307	0.2347

Based on the values in Table 5-8, the differences between the mean values of signals of Cluster 1 and Cluster 2 are obvious. As with the analysis of all turns, the zero value of SWA after normalization is less than 0.5, which shows that when approaching right turns, drivers have more steering to right than left. However, their difference is not that much in 150 frames where the frequency of steering to the right and left are almost the same in 150 frames before right turns.

As before, the zero value of Acceleration after normalization is more than 1 in all three timeframes, which means that approaching right turns, all drivers decelerate. As mentioned before, the higher value in this signal means lower deceleration (smaller negative acceleration). Considering the mean value of the Acceleration signal in Table 5-8, it is obvious that drivers in Cluster1 have lower deceleration than Cluster2 (which have higher deceleration). This means that drivers in Cluster2 reduce their speed more rapidly than those in Cluster1, as a result they push the brake pedal harder and they have lower pressure on the gas pedal.

As the cluster analysis for all turns in the previous section, we can again identify the same types of clusters:

- *Cluster 1 (Moderate Drivers):* Moderate speed, some gas pressure approaching turn, gentle braking, and gradual deceleration;
- *Cluster 2 (Aggressive Drivers):* Higher speed approaching turns, harder braking and more rapid deceleration.

Again, there exists differences between two clusters formed from the 450 frames; this is probably because we are analyzing too many frames and signal values start to level out.

Though not surprising, the analysis of the right turns is very consistent with the clusters identified in the analysis of all turns.

5.3.3 Left Turns

Using statistical features extracted from driving signals before left turns, we perform Hierarchical Clustering Analysis (HCA) on 150, 300, and 450 frames. Figure 5-5 shows the dendogram related to the HCA on 150 frames before left turns; other dendograms related to 300 and 450 frames before left turns is presented in Appendix B.

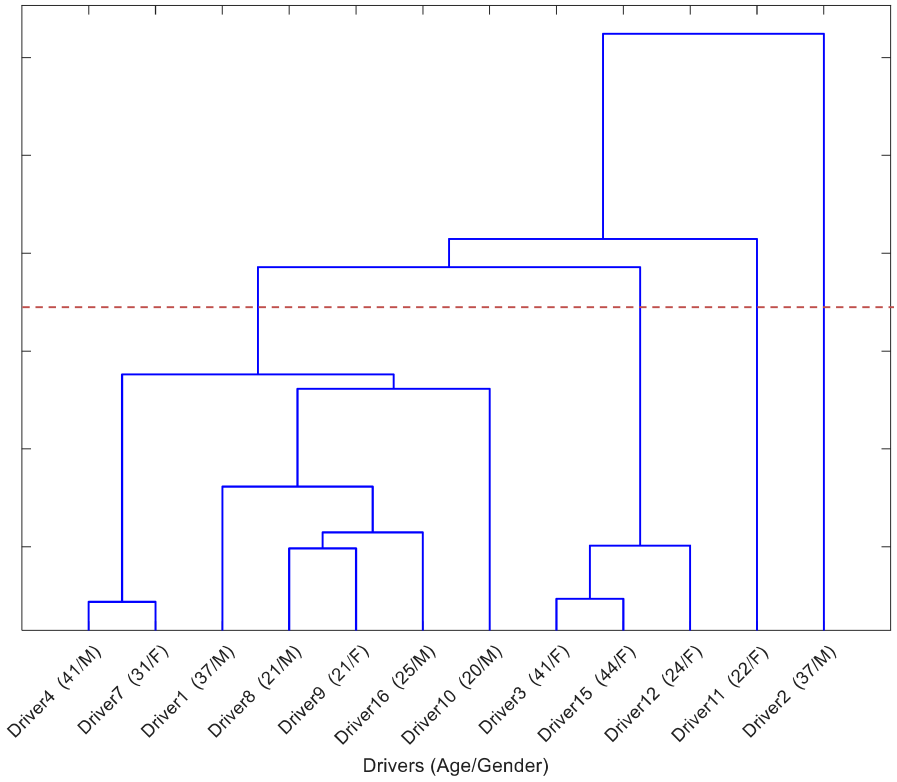

Figure 5-5. HCA based on statistical features considering 150 driving frames before left turns.

We can identify two clusters for each dendograms related to different timeframes before left turns. According to the clustering result shown in Figure 5-5, Cluster 1 includes drivers “1, 10, 8, 9, 16, 4, 7” and Cluster 2 contains drivers “3, 15, 12”. In order to investigate each cluster, we calculate the centroid of each statistical features in each cluster. Table 5-9 shows the centroid values related to HCA on 150 frames before left turns as an instance. Other centroids values related to 300 and 450 frames can be found on Appendix B.

Table 5-9. Centroid values of statistical features in clusters from HCA on 150 driving frames before left turns.

	Cluster 1 "1, 10, 8, 9, 16, 4, 7"					*Cluster 2 "3, 15, 12"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.5841	0.447	0.5606	0.7323	0.5974	0.4549	0.0729	0.6908	0.7037	0.2121
STD	0.0908	0.3786	0.4684	0.0813	0.5472	0.1516	0.1088	0.5541	0.0684	0.4438
Kurtosis	0.0274	0.0943	0.0853	0.1358	0.481	0.0234	0.5919	0.1091	0.2347	0.899
Skewness	0.1873	0.1961	0.2605	0.5017	0.633	0.268	0.8888	0.0708	0.3372	0.5617

Based on the centroid values in Table 5-9, we calculate the Angle and the R-Value between these two vectors same as before. The Angle and R-Value between two clusters in each experiment using 150, 300, and 400 frames are shown in Table 5-10.

Table 5-10. Angle and R-Value between centroid vectors of two clusters result from 150, 300, and 450 frames before left turns.

Left Turns	***Cluster 1 (Drivers No.)***	***Cluster 2 (Drivers No.)***	***Cluster 1 M/F***	***Cluster 2 M/F***	***Angle (degrees)***	***R-Value***
150 Frames	1,4,**7,8,9,10,16**	**3,12**,15	5M, 2F	0M, 3F	35	0.45
300 Frames	**8,9,10**,11,**15,16**	**1,3,4**,7,**12**	3M, 3F	2M, 3F	32	0.48
450 Frames	2,**7,10,15**	**1,3,4**,8,9,11,**12**,16	2M, 2F	4M, 4F	32	0.22

In Table 5-11, the mean value of signals in different timeframes before left turns is presented. The normalized zero value for Steering Wheel and Acceleration for each analysis is also included in this table.

Table 5-11. Mean value of all signals of two clusters result from performing HCA on 150, 300, and 450 frames before left turns.

Left Turns			***Cluster 1***					***Cluster 2***				
	SW = 0	***Acc = 0***	***Mean Speed***	***Mean Gas***	***Mean Brake***	***Mean SW***	***Mean Acc.***	***Mean Speed***	***Mean Gas***	***Mean Brake***	***Mean SW***	***Mean Acc.***
150 Frms	0.087	1.143	0.5841	0.447	0.5606	0.7323	0.5974	0.4549	0.0729	0.6908	0.7037	0.2121
300 Frms	-0.215	1.309	0.4456	0.6525	0.3915	0.6443	0.7774	0.8358	0.3659	0.8512	0.6414	0.2631
450 Frms	-0.475	1.748	0.7154	0.7467	0.4929	0.6598	0.8024	0.5902	0.3705	0.8271	0.3941	0.4180

The values presented in Table 5-11, indicate similar differences between clusters as for all turns and right turns. The zero value of SWA after normalization is way less than 0.5 in 150 frames and negative in the 300 and 450 frames. This indicates that there is very few left steering in the 150, in the 300 and 450 frames before left turns. Also, the zero value for Acceleration after normalization is more than one as for all turns and right turns, which means that we just have deceleration before turns.

Considering the mean value of the Acceleration signal in Table 5-11, it is obvious that drivers in Cluster1 have lower deceleration than those in the Cluster2 (which have higher deceleration). This means that drivers in Cluster2 lower their speed more rapidly than those in Cluster1, and as a result they push the brake pedal harder and they have lower pressure on gas pedal.

As with the previous analysis, we can label the clusters as Moderate and Aggressive drivers for different timeframes prior to left turns:

- *Cluster 1 (Moderate Drivers):* More gas pressure approaching turn, gentle braking, and gradual deceleration
- *Cluster 2 (Aggressive Drivers):* Harder braking and more rapid deceleration

Again, although there exist considerable differences between two clusters in 450 frames, it seems that for 450 frames we are analyzing too much frames because signal values start to level out.

We also looked at some of the cluster analyses where there were 3 clusters, though this was not possible in all situations. As with the 2 cluster situation, we compared the vectors representing the three clusters. Same patterns of drivers show up in the three cluster in those analyses, though the third cluster found in each case seemed to be segmented from one of the pairs of clusters identified in the two cluster situation.

5.4 Cluster analysis using DTW

Dynamic Time Warping (Berndt and Clifford 1994), as mentioned before, is a time-series alignment algorithm, which aims to align two sequences of features by warping the time axis to find an optimal match. Since our data is time-series and we are looking for similarities between driving signals which are not necessarily aligned in a specific period of time, this approach seems to fit our needs.

In order to obtain a better result for comparing the signals, some preprocessing is needed before applying the DTW algorithm. First, we smooth the signals using a smoothing function that is a moving average filter of a specific size. Then we compute the z-score (standard score) for each instance of each signal so that all sequences are centered to have mean equals to 0 and scaled to have standard deviation equals to 1. This will help signals to be aligned vertically and hopefully results in better comparisons.

Using DTW as the distance measure on 150, 300, and 450 frames before turns, we end up with a symmetric12 x 12 distance matrix which indicates the distance between each pair of drivers.

Table 5-12 shows an example of distance matrix based on 300 frames before all turns. Based on this distance matrix, we perform the hierarchical cluster analysis. As with our previous approach, the 'ward' method is used as the linkage criteria for hierarchical cluster analysis.

Table 5-12. An example of a distance matrix results from performing DTW algorithm on all signals 300 frames before all turns.

	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1	0											
D2	74.4932	0										
D3	72.6479	76.1137	0									
D4	68.6884	77.9309	70.8191	0								
D7	87.5164	87.2952	82.4854	86.3546	0							
D8	70.6811	80.6799	82.8156	82.0003	90.5861	0						
D9	71.8016	76.6865	76.1693	80.8976	86.0703	71.2459	0					
D10	75.6432	84.6925	76.0442	78.9084	94.5304	85.5281	85.1184	0				
D11	79.4278	81.7004	76.545	77.0671	93.378	85.4301	81.7125	81.1709	0			
D12	73.4488	85.2893	72.457	81.5628	84.3096	84.1272	76.0266	75.52	71.9166	0		
D15	78.9168	82.2629	74.9728	80.1667	89.0437	82.5796	88.2824	83.4631	85.4309	78.2106	0	
D16	89.577	95.1347	90.7233	91.2382	89.3299	84.3107	87.8455	92.6021	84.5202	90.1001	98.6547	0

In the next three sections, we would study the result of HCA on different timeframes of signals before all turns and also left and right turns separately, using DTW.

5.4.1 All turns

We have applied HCA on 150, 300, and 450 frames before all turns. Figure 5-6 contains the dendogram related to clustering driving behaviour using DTW, based on 300 frames before all turns. Other dendograms regarding 150 and 450 frames can be found in Appendix B.

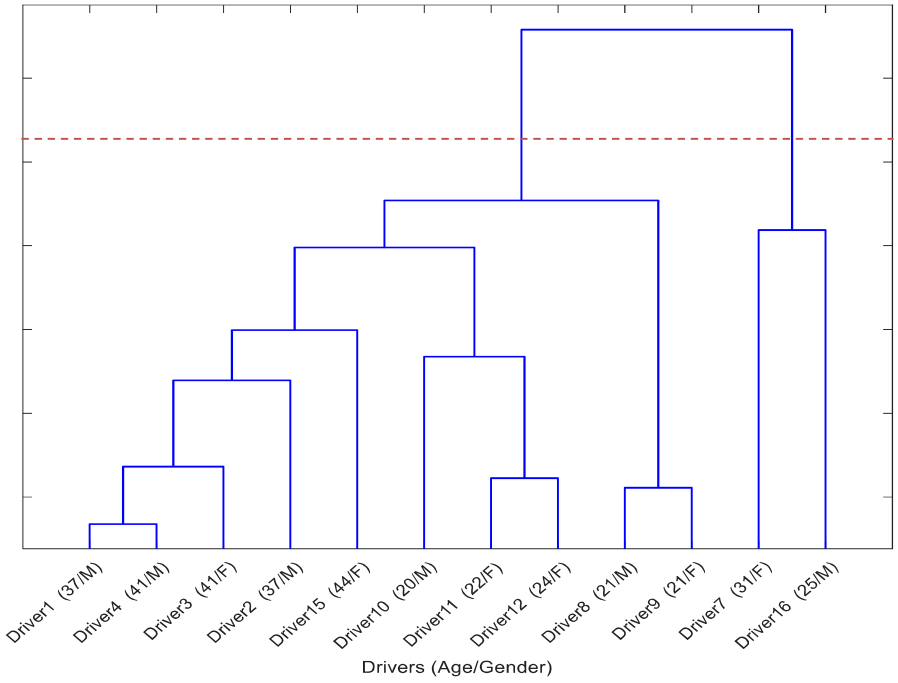

Figure 5-6. HCA based on DTW considering 300 driving frames before all turns.

Table 5-13 presents the drivers in each of the clusters resulting from HCA on different timeframes before all turns using DTW. The results show high consistency between cluster members in different timeframes. In fact, using 150 frames and 450 frames in this case clusters the drivers exactly the same. Since the only parameters that we have in this approach is the information about each driver, namely, gender and age, we are not able to conclude much about the driving characteristics of the drivers in each cluster since both clusters have both male and female drivers in various age ranges.

Table 5-13. Summary result of clustering driver behaviour using DTW with different timeframes before all turns.

	Cluster 1 (Drivers No.)	*Cluster 2 (Drivers No.)*
150 Frames	**1,2,3,4**,7,**10,11,12,15**	**8,9,16**
300 Frames	**1,2,3,4**,8,9,**10,11,12,15**	7,**16**
450 Frames	**1,2,3,4**,7,**10,11,12,15**	**8,9,16**

In order to investigate if there exist any differences in driving behaviour characteristics approaching to right or left turns, cluster analysis is performed on right turns and left turns separately. The result is presented in the next two sub-sections.

5.4.2 Right Turns

As with the previous sub-section we apply HCA on 150, 300, and 450 frames before right turns. The dendogram related to applying HCA on 300 frames prior to right turns using DTW is shown in Figure 5-7. Other dendograms related to similar cluster analysis on 150 and 450 frames can be found in Appendix B.

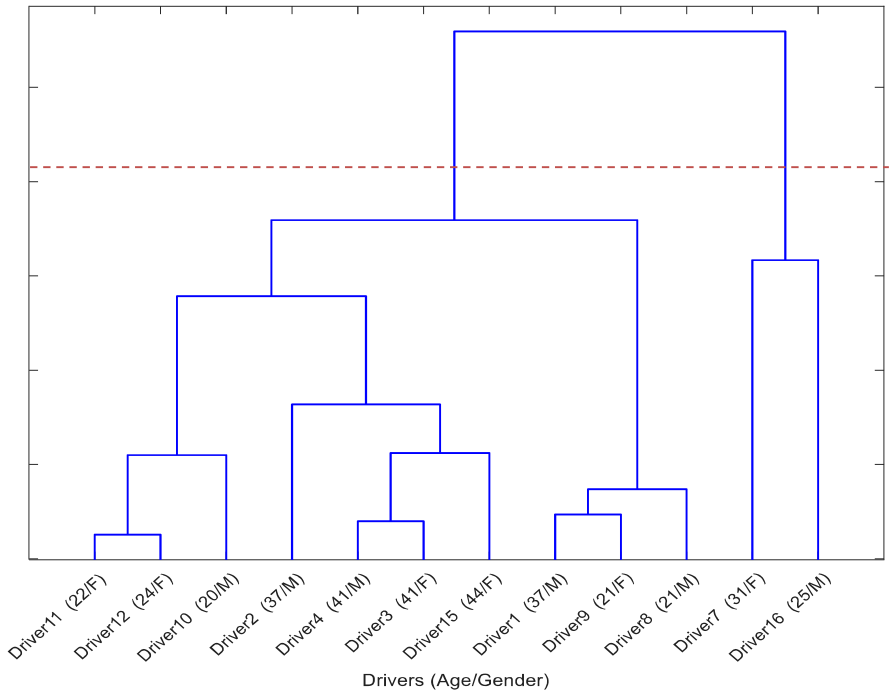

Figure 5-7. HCA based on DTW considering 300 driving frames before right turns.

Table 5-14 contains drivers' number in each clusters result from HCA on different timeframes prior to right turns using DTW.

Table 5-14. Summary result of clustering driver behaviour using DTW with different timeframes before right turns.

	Cluster 1 (Drivers No.)	***Cluster 2 (Drivers No.)***
150 Frames	**1,2,3,4**,7,**10,11,12,15**	8,9,**16**
300 Frames	**1,2,3,4**,8,9,**10,11,12,15**	**7,16**
450 Frames	**1,2,3,4**,8,9,**10,11,12,15**	**7,16**

The results show high consistency between cluster members in different timeframes. In fact, HCA analysis using 300 frames and 450 frames results in the same clusters as the HCA analysis with 150 frames. Once again, the results for right turns analysis is almost the same as all turn analysis. Again, we can conclude little about the clusters.

5.4.3 Left Turns

Using the DTW technique as a distance measure we perform the hierarchical clustering for pre-left-turn driving behaviour. As with the previous analyses, we study 150, 300, and 450 non-zero-speed frames before left turns. Dendogram related to HCA result on 300 frames is shown in Figure 5-8. Other dendograms for 150 and 450 frames can be found in Appendix B.

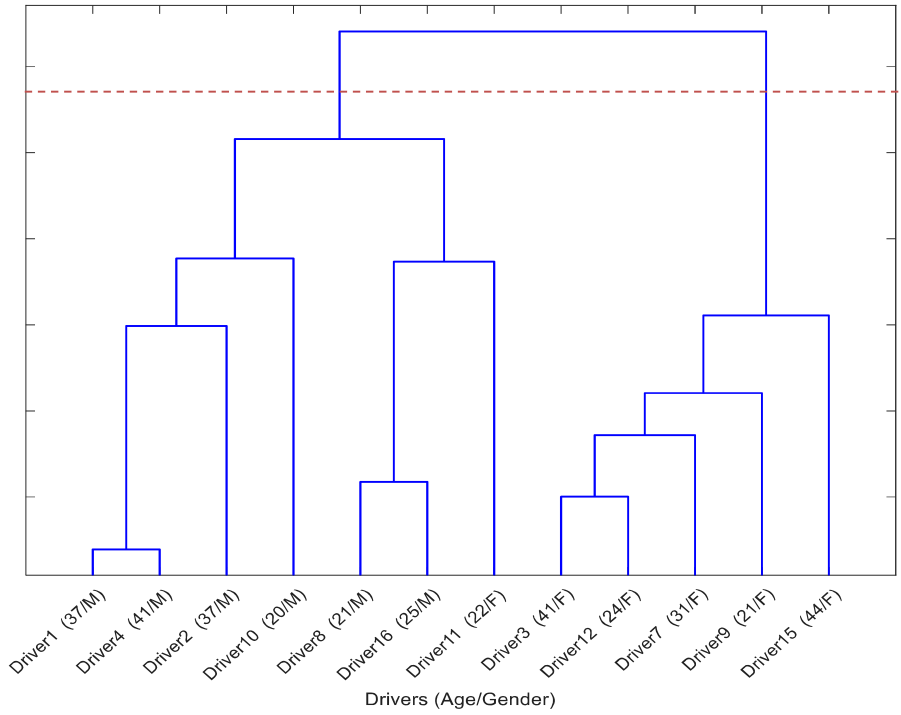

Figure 5-8. Hierarchical clustering analysis based on DTW considering 300 driving frames before left turns.

The result of the cluster analysis performed on left turns with different timeframes is presented in Table 5-15.

Table 5-15. Summary result of clustering driver behaviour using DTW with different timeframes before right turns.

	Cluster 1 (Drivers No.)	***Cluster 2 (Drivers No.)***
150 Frames	**1,2,3,4,7,10,12,15**	**8,**9,**11**,**16**
300 Frames	**3,7,9,12,15**	1,2,4,**8,**10,**11,16**
450 Frames	**1,2,3,4,7,9,10, 12,15**	**8,11,16**

Similar to the previous result, this analysis also reflected consistency of members across clusters. Moreover, if we look at the analysis for 300 frames, we see a clustering along gender lines. As shown in Figure 5-8, we can see that male and female drivers are grouped predominantly into two clusters. As shown in Table 5-16, based on this clustering analysis, 100% of the drivers in Cluster 1 are female drivers and 85.7% of the drivers in Cluster 2 are male drivers. This suggests that there may be some gender

differences in how they approach left turn; however, these results would need to be validated on a larger dataset with greater numbers of drivers.

Table 5-16. Two clusters result from cluster analysis using DTW for 300 pre-left-turns driving frames.

Cluster	Drivers	Ages	Genders	Percent
1	3, 12, 7, 9, 15	41, 24, 31, 21, 44	F, F, F, F, F	% 100
2	1, 4, 2, 10, 8, 16, 11	37, 41, 37, 20, 21, 25, 22	M, M, M, M, M, M, F	% 85.7

5.5 Discussion of Overall Results

Considering the results from previous sections, we show that analyzing the statistical features of CAN-bus signals would result in at least two distinct clusters of driving behaviour before turns, indicating moderate (normal) or aggressive driving behaviour. Moreover, the consistency between clusters' members was good throughout different timeframes and even for left and right turns separately.

The results of our second approach in clustering drivers' behaviour using DTW indicate that the consistency between the membership of the clusters member was high across different timeframes and there was even some consistency across the clusters obtained from all turns, left turns and right turns. In the cluster analysis on 300 frames using the DTW approach before left turns (Figure 5-8), male and female drivers seem to be categorized into two different groups suggesting that there may be some gender differences in driving behaviour approaching a left turn. This result may have occurred because of left turns seem to be more challenging when compared to right turns and may result in different behaviour. This is, however, a single example and more data is needed to investigate this hypothesis.

Chapter 6

6 Conclusion and Future Works

Understanding driver behaviour can be used to improve Advanced Driver Assistance Systems (ADASs), improve vehicle safety and privacy by monitoring the driver's behaviour, and also help to detect risky driving styles. Driving behaviour has been studied from many aspects in the field of traffic safety analysis. Many of these studies have focused on various aspects of overall driving behaviour of individual drivers and have met with varying success. Few of these studies focus on specific parts of driving to analyze individual driving characteristics. Driving activities in preparing for a turn is a complex driving behavior. It is a confined period of activities where there would appear to be identifiable differences in driver behavior. We have been able to identify these differences in small time periods (5-15 seconds) before turns. Analysis of driving behaviour in different maneuvers may enable the intelligent driver assistance systems to be customized for individual drivers.

6.1 Conclusion

In this thesis, driver behaviour is analyzed in different timeframes prior to turns. Our aim was to study the way each driver would prepare the vehicle for a turning maneuver and to find possible distinct clusters that represent their behaviour. We carried out the investigation on actual driving data collected from 12 drivers driving through a pre-determined path in an urban area inside the city of London, Ontario. Five CAN-Bus time-series signals, including speed, gas pedal pressure, brake pedal pressure, steering wheel angle, and acceleration, were collected for 5, 10, and 15 seconds before all turns; left turns and right turns were also considered separately.

We applied two different approaches to cluster drivers using these data. In the first approach, 4 statistical features including *Mean*, *Standard Deviation*, *Kurtosis*, and *Skewness* were extracted from the signals. Using these statistical features as a representative of each driver's behaviour, Hierarchical Clustering Analysis was used to

cluster the drivers. The results show that there exist at least two distinct groups of drivers with different pre-turn behaviour, one cluster includes *moderate* drivers, while the other cluster contains more *aggressive* ones.

Another approach carried out in this study was using Dynamic Time Warping (DTW) to measure the dissimilarity between different time-series signals of drivers. As the previous approach, we used Hierarchical Clustering Analysis to cluster the drivers based on the dissimilarity matrix we calculated through DTW. The results show high consistency between members of clusters in different timeframes. Also, in the case of 300 frames before left turns, male and female drivers categorized into two different groups.

A major distinction of our approach was to focus on a small portion of a driving behaviour i.e. before turns, and cluster drivers based on that. At least two distinct clusters of driving behaviour detected by our analysis.

6.2 Future Work

This novel approach has great potential in several fields. Future work will focus on the testing our approach on larger dataset, containing more drivers and more turns. Also, we would like to see if it is possible to use this approach to classify drivers, by collecting and computing data while driving and then mapping a driver to a class. Further, we could use the cluster characteristics identified and investigate whether it might be possible to determine if a driver was not preparing for a turn by comparing the immediate data to the cluster corresponding to that driver. This might provide a method of providing an “early warning” when a driver is not paying attention to turns.

Appendix A

Summary of Statistical Features for All Drivers Over All Turns

Table A-1. Statistical features for each signal for all drivers over all turns.

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 1	Speed	33.097	9.7282	2.5267	-0.5157
	Gas	10.49	16.0562	5.0206	1.6615
	Brake	42.4679	38.2385	2.1391	0.178
	SWA	4.0627	12.4934	3.2634	0.0477
	Acceleration	-1.1597	3.2778	2.9281	0.0513

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 2	Speed	29.6391	12.45	17.7962	0.7574
	Gas	11.0025	13.8383	4.8364	1.2215
	Brake	41.7044	38.0499	2.1371	0.3014
	SWA	-0.1947	19.4537	28.4443	0.9429
	Acceleration	-0.8067	2.7939	2.6563	-0.3728

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 3	Speed	34.2961	8.7619	2.65	-0.6865
	Gas	7.8214	11.5761	7.6512	1.681
	Brake	39.0537	31.363	2.5021	0.3674
	SWA	3.221	12.748	3.2795	0.1046
	Acceleration	-1.5882	2.2828	2.2762	-0.3111

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 4	Speed	34.4621	9.1342	2.5879	-0.5702
	Gas	11.7788	15.7949	9.1431	1.8641
	Brake	41.1542	34.2111	2.8136	0.1091
	SWA	4.7638	10.8893	3.9507	0.1289
	Acceleration	-1.2395	2.9485	2.913	-0.0337

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 7	Speed	28.0978	8.7935	2.5904	-0.29
	Gas	12.4142	15.8853	2.4068	0.8369
	Brake	35.4999	37.3366	1.7302	0.3681
	SWA	5.9341	20.3927	3.9916	0.4325
	Acceleration	-0.6849	3.0551	2.1369	-0.1353

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 8	Speed	24.0304	7.5711	2.5058	-0.2098
	Gas	7.7779	11.1518	4.9558	1.4225
	Brake	39.9979	35.2268	2.461	0.2908
	SWA	1.3371	13.5558	3.1047	0.0012
	Acceleration	-0.9034	2.4741	2.297	-0.1105

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 9	Speed	26.7962	9.105	2.6107	-0.1643
	Gas	6.8268	9.7566	7.6564	1.539
	Brake	43.7809	28.9318	2.7092	-0.3379
	SWA	7.2792	8.3119	4.215	0.4574
	Acceleration	-1.3697	2.4579	2.8831	-0.209

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 10	Speed	29.1785	8.2703	2.782	-0.4831
	Gas	12.8152	13.8042	3.4816	0.8659
	Brake	24.8244	30.6001	3.1584	0.9766
	SWA	5.4873	17.693	3.4838	0.1549
	Acceleration	-0.4034	2.6566	2.882	-0.5161

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 11	Speed	24.9454	8.948	2.973	-0.1568
	Gas	6.4433	7.0318	7.2759	1.5657
	Brake	40.7504	28.6675	5.2453	0.4365
	SWA	-0.7094	12.6119	4.102	-0.1083
	Acceleration	-1.2563	2.5818	2.6246	-0.5

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 12	Speed	32.1247	10.813	2.3307	-0.4055
	Gas	12.9153	16.8514	3.0522	0.9894
	Brake	41.127	34.8289	2.1598	0.1025
	SWA	2.8904	12.8679	3.9028	-0.0096
	Acceleration	-1.2437	3.0067	2.0755	-0.0978

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 15	Speed	29.3951	9.499	2.6379	-0.3099
	Gas	12.0071	15.0025	4.8097	1.3208
	Brake	36.2827	36.3882	2.2696	0.3644
	SWA	3.0584	18.9255	3.5816	0.0713
	Acceleration	-1.0546	3.0777	2.1584	-0.2291

		Mean	*STD*	*Kurtosis*	*Skewness*
Driver 16	Speed	21.4632	8.6786	2.2643	-0.1088
	Gas	7.8826	12.5062	4.3079	1.4086
	Brake	45.5087	32.6084	2.023	-0.1293
	SWA	5.757	11.6498	3.4747	0.111
	Acceleration	-1.1218	3.1164	2.4729	0.0818

Appendix B

Summary of Cluster Analyses and Centroids

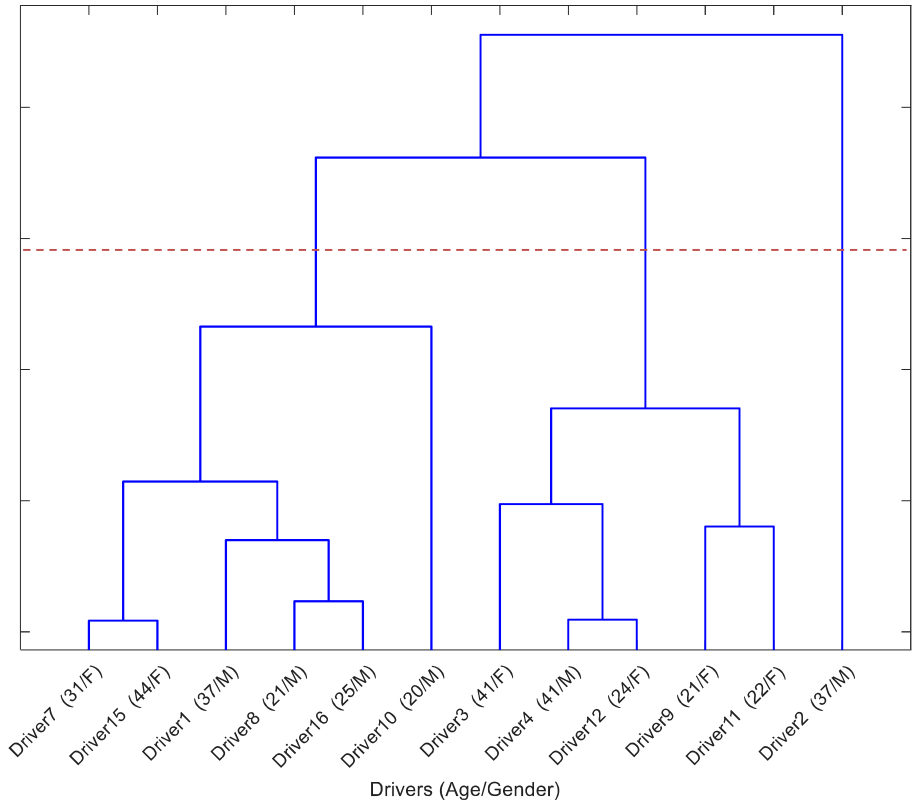

Figure B-1. HCA based on Statistical Features on 300 driving frames before all turns.

Table B-1. Centroid values of statistical features in clusters from HCA on 300 driving frames before all turns.

	Cluster 1 *"7, 15, 1, 8, 16, 10"*					*Cluster 2* *"3,4,12, 9, 11"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.5445	0.7271	0.5316	0.5588	0.7914	0.694	0.3181	0.8608	0.6107	0.3103
STD	0.1515	0.7494	0.6775	0.3423	0.6866	0.2097	0.2663	0.2408	0.1412	0.2644
Kurtosis	0.0437	0.1239	0.1615	0.0609	0.2958	0.0484	0.169	0.5374	0.0954	0.4806
Skewness	0.2054	0.4212	0.5575	0.4963	0.6278	0.1681	0.4972	0.3342	0.3469	0.5313

Zero-SWA = 0.1928

Zero-ACC = 1.7536

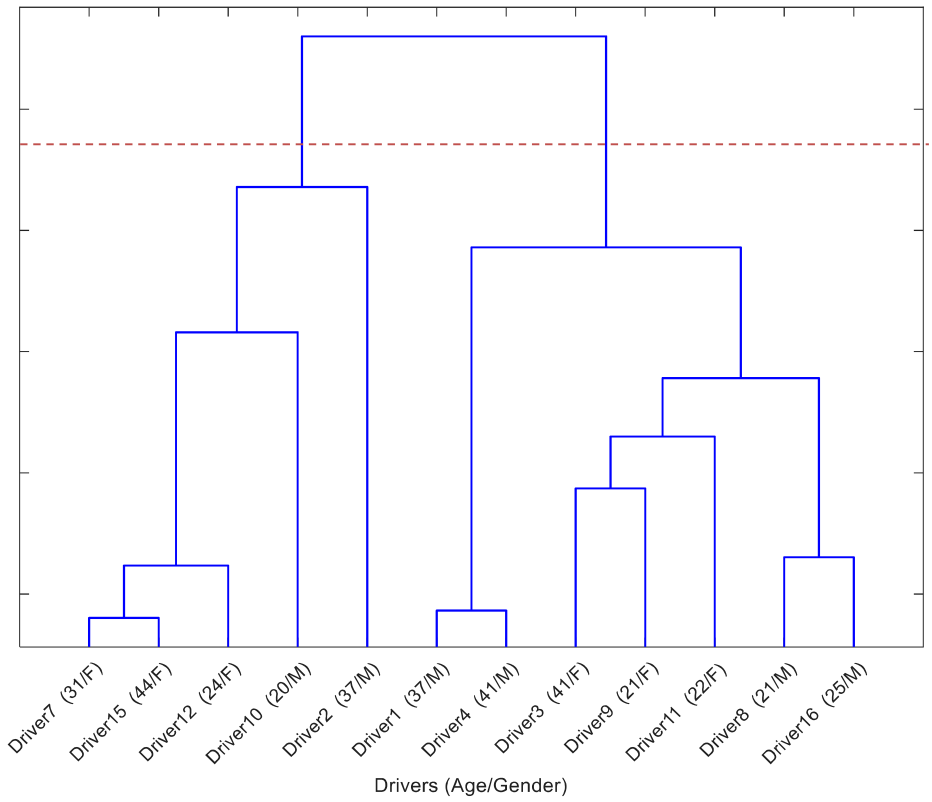

Figure B-2. HCA based on Statistical Features on 450 driving frames before all turns.

Table B-2. Centroid values of statistical features in clusters from HCA on 450 driving frames before all turns.

	Cluster 1 *"7, 15, 12, 10, 2"*					*Cluster 2* *"1, 4, 3, 9, 11, 8, 16"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.6327	0.8942	0.5349	0.5188	0.6326	0.5368	0.3072	0.8215	0.5486	0.2989
STD	0.4907	0.8192	0.7077	0.7909	0.6383	0.2615	0.5041	0.4265	0.2847	0.4536
Kurtosis	0.2165	0.1945	0.1595	0.2201	0.3592	0.0209	0.6185	0.3162	0.0206	0.6477
Skewness	0.3742	0.2044	0.5785	0.4059	0.4112	0.2368	0.7348	0.3564	0.2039	0.6167

Zero-SWA = 0.0888

Zero-ACC = 1.3403

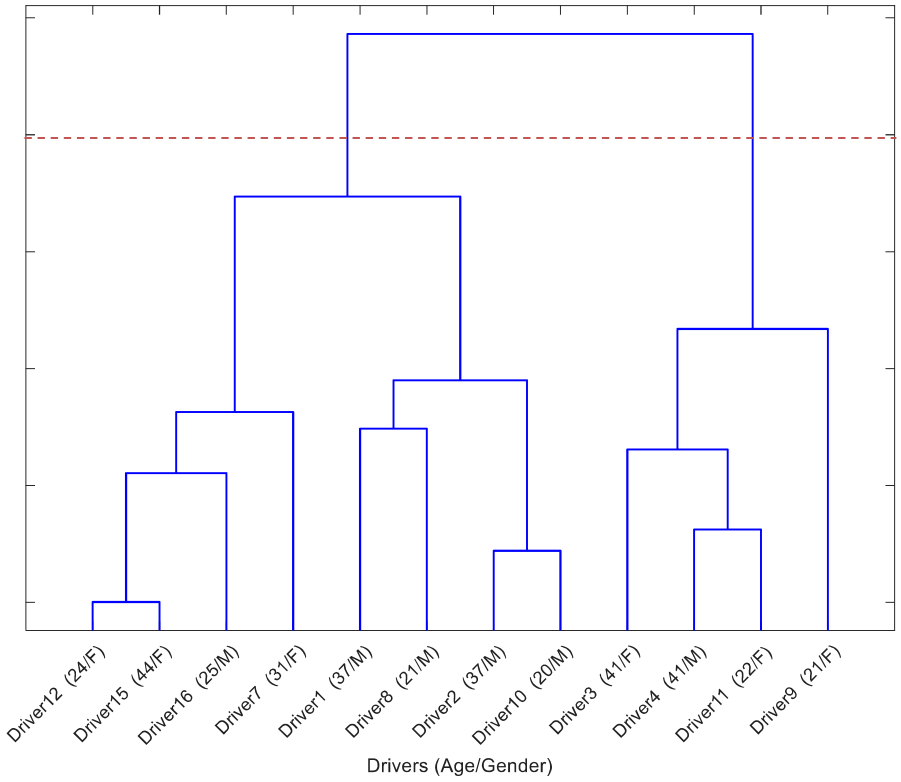

Figure B-3. HCA based on Statistical Features on 300 driving frames before right turns.

Table B-3. Centroid values of statistical features in clusters from HCA on 300 driving frames before right turns.

	Cluster 1 *"12, 15, 16, 7,1,8,2,10"*					*Cluster 2* *"3, 4, 11, 9"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.5384	0.699	0.4022	0.5162	0.7393	0.6814	0.2368	0.8092	0.7114	0.1589
STD	0.6278	0.5722	0.6311	0.4712	0.5448	0.6549	0.1121	0.1352	0.1142	0.3657
Kurtosis	0.3694	0.2108	0.1428	0.3079	0.3242	0.8068	0.2362	0.5511	0.757	0.5426
Skewness	0.7061	0.388	0.7331	0.3919	0.6548	0.3579	0.3649	0.6129	0.3527	0.3923

Zero-SWA = 0.3700

Zero-ACC = 1.4805

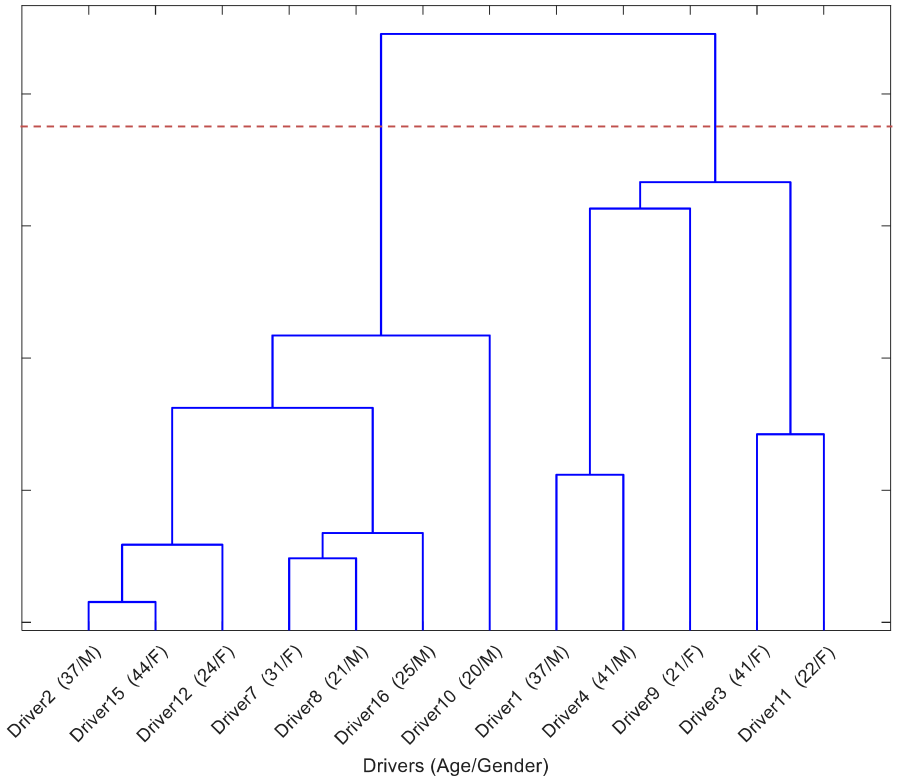

Figure B-4. HCA based on Statistical Features on 450 driving frames before right turns.

Table B-4. Centroid values of statistical features in clusters from HCA on 450 driving frames before right turns.

	Cluster 1 *"12, 15, 2, 8, 16, 7, 10"*					*Cluster 2* *"1, 4, 3, 11, 9"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.391	0.6713	0.5391	0.4375	0.6037	0.704	0.4322	0.7309	0.5307	0.2347
STD	0.6313	0.5786	0.6651	0.6366	0.5606	0.6255	0.3896	0.393	0.2057	0.4992
Kurtosis	0.2726	0.1444	0.1059	0.2945	0.2473	0.8184	0.5208	0.3404	0.433	0.681
Skewness	0.7756	0.2273	0.5979	0.2925	0.675	0.2636	0.6042	0.5536	0.4275	0.5254

Zero-SWA = 0.2565

Zero-ACC = 1.0368

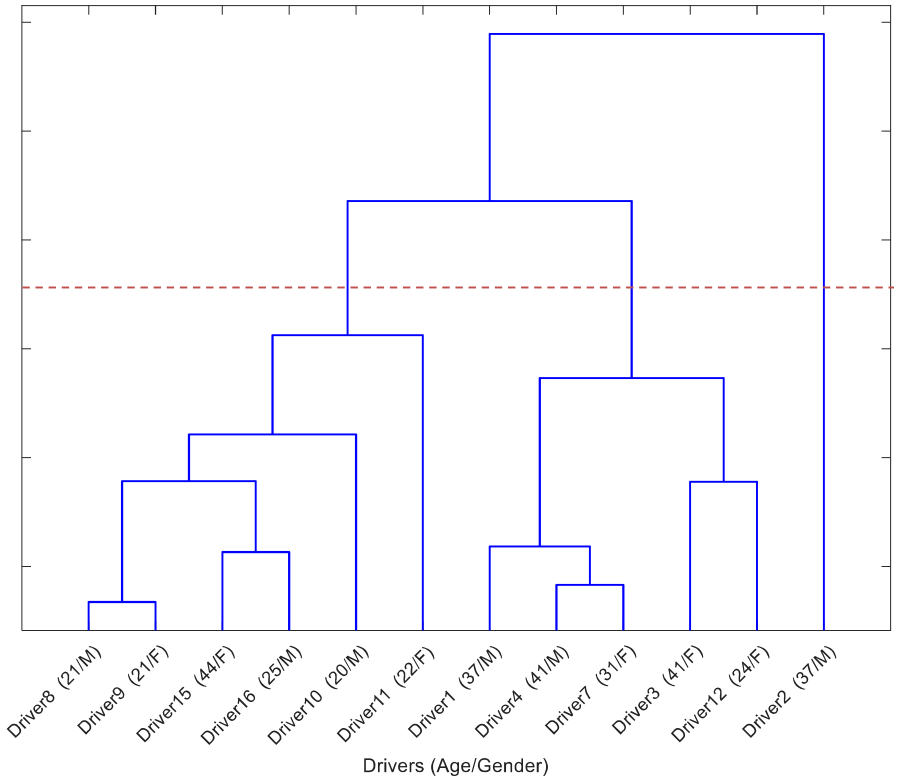

Figure B-5. HCA based on Statistical Features on 300 driving frames before left turns.

Table B-5. Centroid values of statistical features in clusters from HCA on 300 driving frames before left turns.

	Cluster 1 *"8, 9, 11, 10, 15, 16"*					*Cluster 2* *"1, 4, 7, 3, 12"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.4456	0.6525	0.3915	0.6443	0.7774	0.8358	0.3659	0.8512	0.6414	0.2631
STD	0.1256	0.571	0.4832	0.052	0.4734	0.3215	0.531	0.6321	0.1894	0.6596
Kurtosis	0.0221	0.0572	0.2881	0.0762	0.1628	0.0169	0.1105	0.4319	0.0413	0.6371
Skewness	0.22	0.1503	0.5241	0.312	0.4447	0.2496	0.2932	0.1821	0.3131	0.7642

Zero-SWA = -0.2153

Zero-ACC = 1.3087

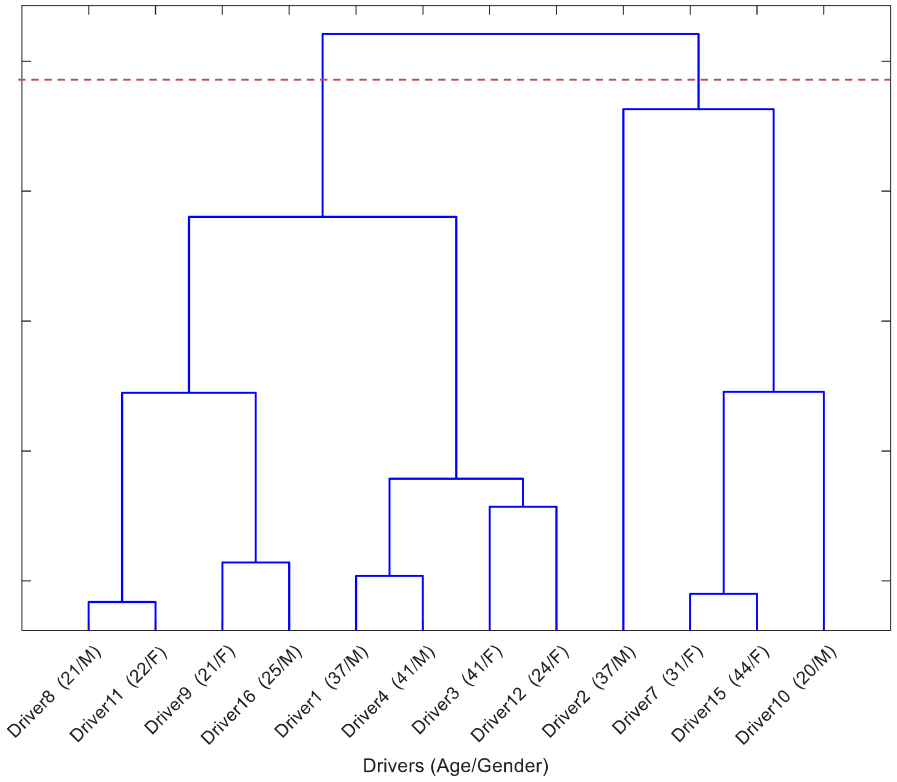

Figure B-6. HCA based on Statistical Features on 450 driving frames before left turns.

Table B-6. Centroid values of statistical features in clusters from HCA on 450 driving frames before left turns.

	Cluster 1 *"8, 11, 9, 16, 1, 4, 3, 12"*					*Cluster 2* *"2, 7, 15, 10"*				
	Speed	Gas	Brake	Steer	Acc	Speed	Gas	Brake	Steer	Acc
Mean	0.5902	0.3705	0.8271	0.3941	0.418	0.7154	0.7467	0.4929	0.6598	0.8024
STD	0.234	0.5258	0.4317	0.1691	0.3909	0.3464	0.8802	0.731	0.5612	0.7957
Kurtosis	0.0136	0.4419	0.3151	0.0161	0.461	0.2757	0.1401	0.3856	0.2655	0.4586
Skewness	0.1897	0.6264	0.1638	0.1594	0.604	0.3422	0.2638	0.5835	0.4882	0.2742

Zero-SWA = -0.4745

Zero-ACC = 1.7478

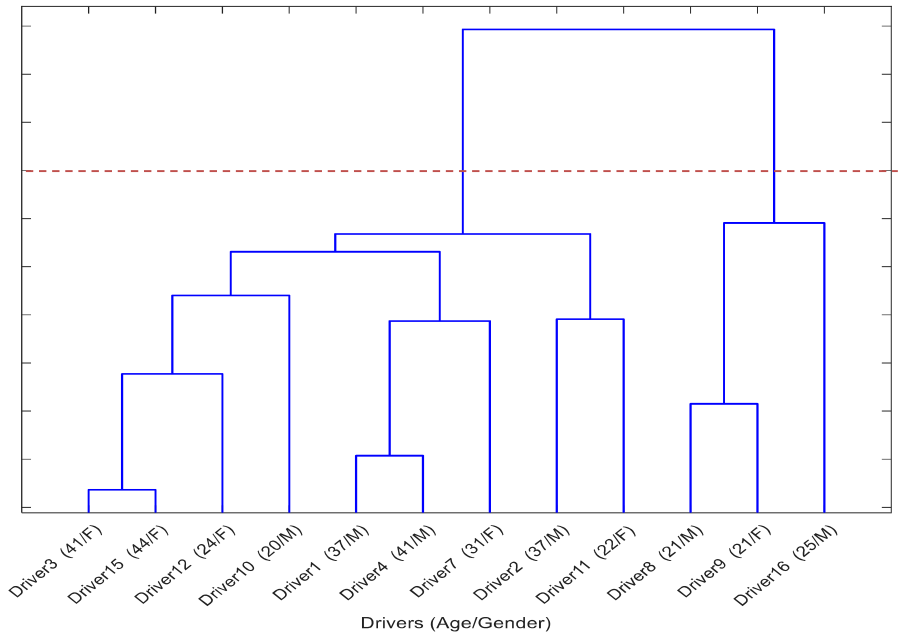

Figure B-7. HCA based on DTW considering 150 driving frames before all turns.

Table B-7. Distance matrix results from performing DTW algorithm on all signals 150 frames before all turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	51.2546											
D3	48.6483	49.5566										
D4	44.3311	52.4798	46.4935									
D7	51.5046	54.6996	51.348	49.7418								
D8	48.308	55.1185	54.0396	54.1882	60.0213							
D9	50.5568	53.9396	52.0281	53.8165	54.7563	47.5004						
D10	52.1599	57.877	49.1979	54.0064	59.5241	59.2137	58.8579					
D11	54.2962	52.671	48.2999	51.4589	58.6577	57.2593	55.0249	56.2285				
D12	51.7494	56.1508	45.3792	50.6812	50.2555	54.2545	55.8975	52.7679	50.7702			
D15	48.6052	50.7403	42.2493	46.9675	52.734	48.872	55.5972	52.5592	53.6186	49.8345		
D16	57.9264	68.3938	59.5456	59.0409	58.5393	53.9675	57.9566	63.5366	58.9484	59.1714	62.6138	

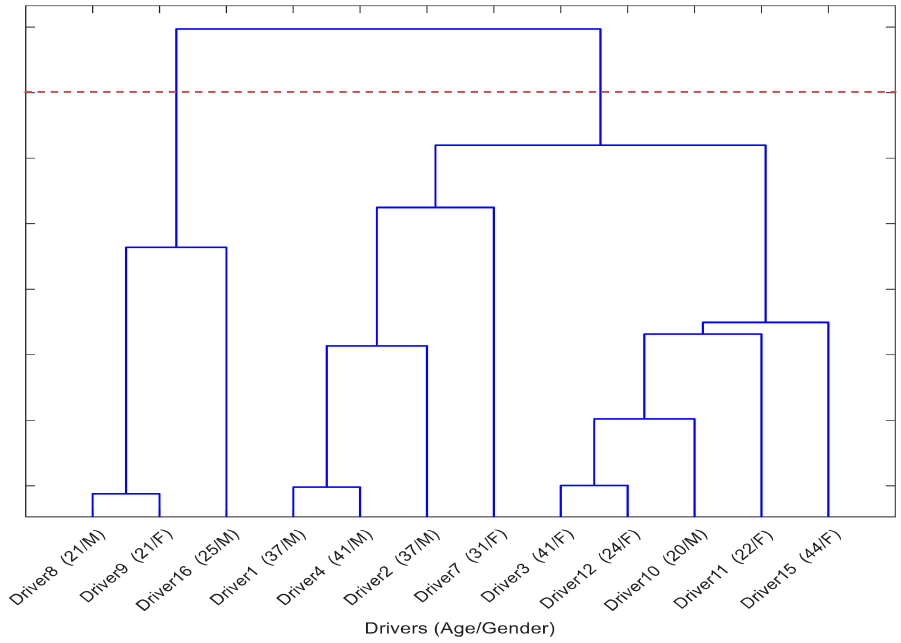

Figure B-8. HCA based on DTW considering 450 driving frames before all turns.

Table B-8. Distance matrix results from performing DTW algorithm on all signals 450 frames before all turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	94.9468											
D3	93.1676	101.148										
D4	88.1167	100.541	94.3533									
D7	106.599	108.045	100.058	107.526								
D8	97.7012	102.060	102.294	109.326	114.368							
D9	95.0492	99.1176	94.4158	108.968	108.005	87.5281						
D10	98.3847	108.921	91.5471	102.881	109.686	105.695	105.672					
D11	106.717	104.767	99.1274	104.075	119.979	108.655	102.161	99.6747				
D12	93.9888	104.766	88.2726	107.591	102.878	105.627	96.6164	93.9993	96.5483			
D15	105.698	102.099	95.5567	102.434	109.632	103.964	108.075	104.288	104.026	95.7589		
D16	113.924	117.946	117.577	115.291	110.427	100.546	108.403	113.165	106.025	114.860	121.426	

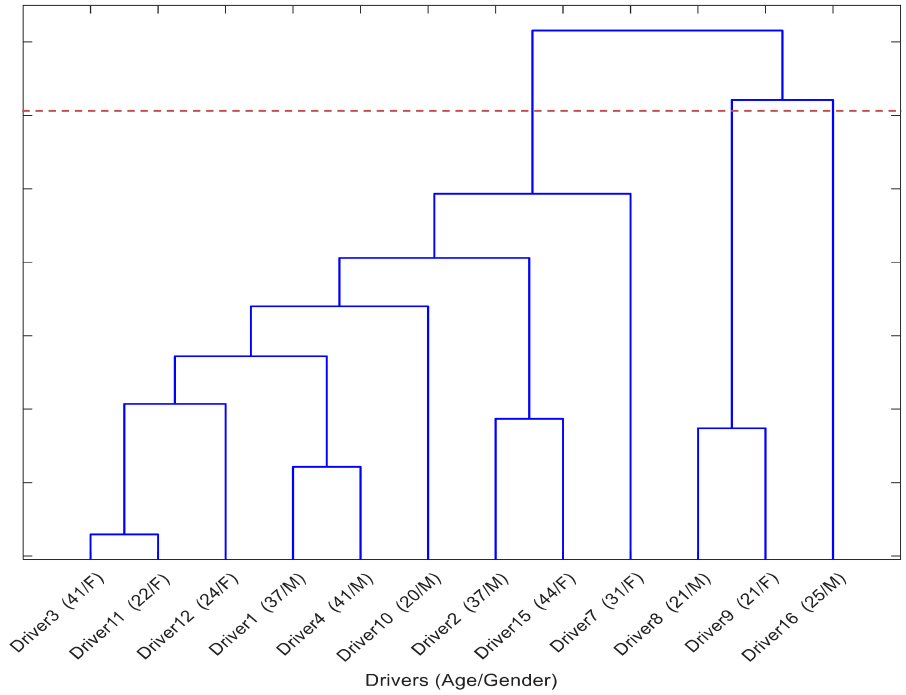

Figure B-9. HCA based on DTW considering 150 driving frames before right turns.

Table B-9. Distance matrix results from performing DTW algorithm on all signals 150 frames before right turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	49.5135											
D3	43.3185	44.6568										
D4	40.9993	50.6249	39.8005									
D7	54.6068	54.4743	50.7884	50.2951								
D8	46.1774	51.7396	50.4646	53.6885	63.5316							
D9	48.0922	51.1961	50.8718	53.1112	55.3085	43.0078						
D10	49.2868	54.7977	45.3592	50.5047	59.8392	62.4929	59.3831					
D11	44.5927	44.4074	36.2492	44.445	55.5378	52.5564	46.27	47.7645				
D12	47.8186	56.3619	41.6595	44.8404	51.8492	54.2649	53.7997	49.3847	44.7228			
D15	49.046	43.8538	40.5371	41.5657	56.1563	49.3988	55.9017	46.3685	48.5281	50.3906		
D16	61.9111	71.8039	56.2177	62.9226	60.6855	59.5236	63.0579	61.7722	59.7957	52.2851	67.948	

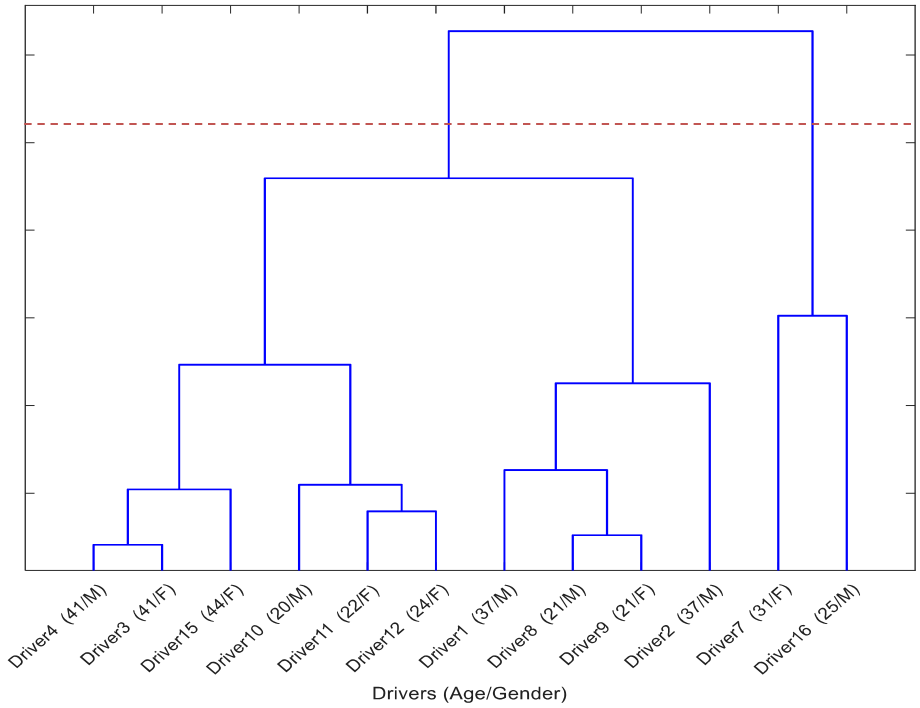

Figure B-10. HCA based on DTW considering 450 driving frames before right turns.

Table B-10. Distance matrix results from performing DTW algorithm on all signals 450 frames before right turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	98.2756											
D3	90.266	98.3832										
D4	90.261	99.4485	83.831									
D7	117.081	110.940	109.438	112.633								
D8	93.5974	99.8389	99.938	104.937	128.205							
D9	87.3721	95.9247	94.5126	102.377	120.085	84.5878						
D10	102.119	107.316	88.3997	94.9056	116.373	107.572	97.5258					
D11	99.8555	94.2353	86.7186	92	115.005	105.693	89.9617	89.7028				
D12	91.858	104.964	94.7665	101.558	111.717	109.945	100.990	90.1775	87.7965			
D15	103.302	95.8731	90.2139	87.0101	113.616	104.884	109.288	96.8859	93.6847	96.0448		
D16	123.409	122.877	114.787	117.873	110.107	110.285	113.590	111.361	107.402	112.456	120.039	

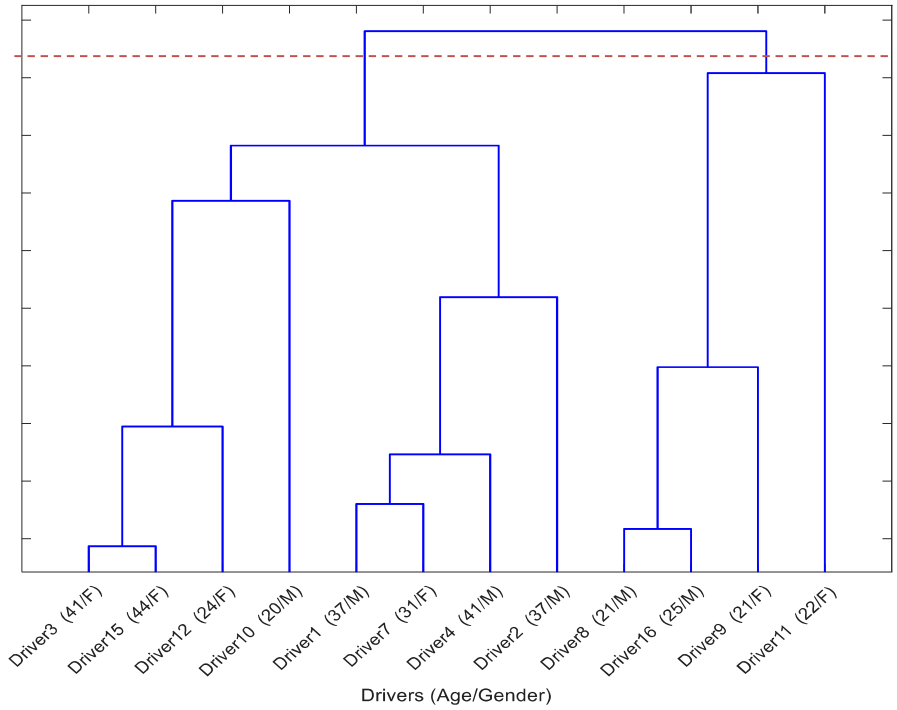

Figure B-11. HCA based on DTW considering 150 driving frames before left turns.

Table B-11. Distance matrix results from performing DTW algorithm on all signals 150 frames before left turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	53.7418											
D3	56.2622	56.5563										
D4	49.0908	55.1296	56.0549									
D7	47.0728	55.0214	52.1473	48.9514								
D8	51.3517	59.9456	59.1468	54.902	55.0064							
D9	54.0777	57.8588	53.6799	54.8242	53.9674	53.9185						
D10	56.2643	62.2758	54.6819	59.009	59.0739	54.5291	58.1076					
D11	68.1583	64.4762	65.5152	61.4788	63.1146	63.9779	67.5318	68.3201				
D12	57.3648	55.8492	50.6932	59.0251	47.9789	54.2397	58.8944	57.6011	59.4093			
D15	47.9754	60.5781	44.6952	54.6845	47.8451	48.1195	55.1622	61.4029	60.8908	49.0401		
D16	52.2339	63.5223	64.2997	53.4956	55.4732	46.0302	50.6691	66.0573	57.738	69.0091	54.9936	

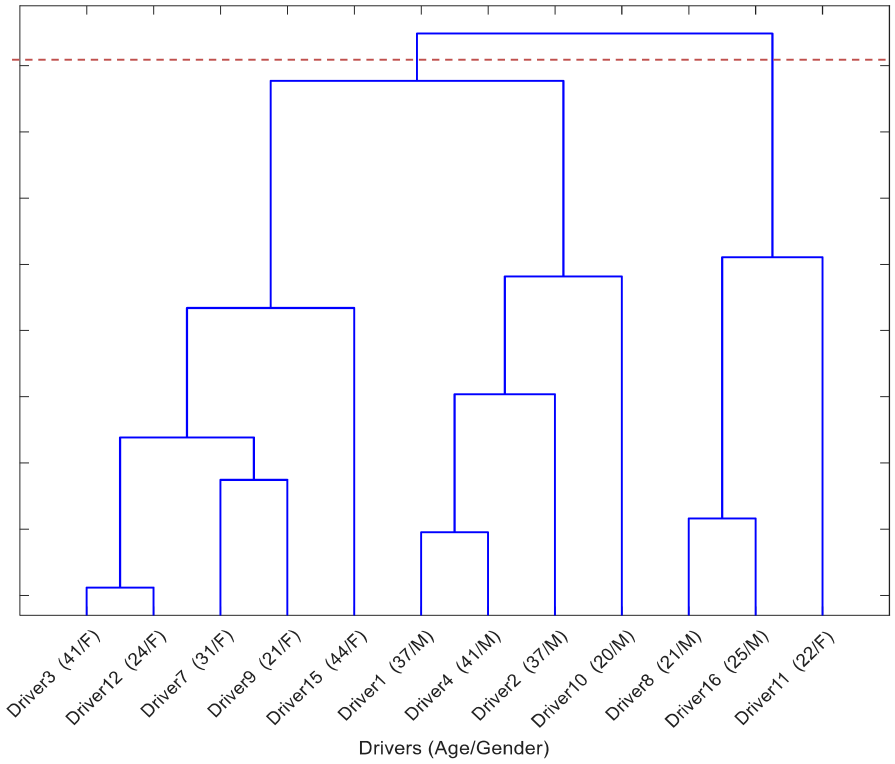

Figure B-12. HCA based on DTW considering 450 driving frames before left turns.

Table B-12. Distance matrix results from performing DTW algorithm on all signals 450 frames before left turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	90.1913											
D3	97.3126	105.098										
D4	85.0535	102.101	109.385									
D7	91.6242	103.908	86.6582	100.232								
D8	103.563	105.234	105.661	115.596	94.6008							
D9	106.016	103.678	94.2776	118.384	90.7497	91.7284						
D10	93.0497	111.213	96.0433	114.275	100.134	103.013	117.310					
D11	116.520	119.812	116.854	121.326	127.086	112.886	119.589	113.920				
D12	97.0326	104.483	78.9956	116.211	90.251	99.4583	90.3677	99.4588	109.051			
D15	109.120	110.994	103.189	124.469	103.941	102.649	106.342	114.863	118.800	95.3504		
D16	100.375	110.902	121.562	111.604	110.883	86.6324	100.992	115.742	104.057	118.295	123.406	

Table B-13. An example of a distance matrix results from performing DTW algorithm on all signals 300 frames before all turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1	0											
D2	74.4932	0										
D3	72.6479	76.1137	0									
D4	68.6884	77.9309	70.8191	0								
D7	87.5164	87.2952	82.4854	86.3546	0							
D8	70.6811	80.6799	82.8156	82.0003	90.5861	0						
D9	71.8016	76.6865	76.1693	80.8976	86.0703	71.2459	0					
D10	75.6432	84.6925	76.0442	78.9084	94.5304	85.5281	85.1184	0				
D11	79.4278	81.7004	76.545	77.0671	93.378	85.4301	81.7125	81.1709	0			
D12	73.4488	85.2893	72.457	81.5628	84.3096	84.1272	76.0266	75.52	71.9166	0		
D15	78.9168	82.2629	74.9728	80.1667	89.0437	82.5796	88.2824	83.4631	85.4309	78.2106	0	
D16	89.577	95.1347	90.7233	91.2382	89.3299	84.3107	87.8455	92.6021	84.5202	90.1001	98.6547	0

Table B-14. An example of a distance matrix results from performing DTW algorithm on all signals 300 frames before right turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	72.7475											
D3	67.2147	72.1581										
D4	70.1736	73.1471	63.7107									
D7	94.4698	86.8538	86.9808	90.1276								
D8	67.2627	77.2001	78.2481	80.3248	100.547							
D9	64.2839	71.753	73.5898	78.1092	92.6175	65.1311						
D10	72.6604	81.8688	71.6455	70.7454	98.1962	86.8711	76.707					
D11	71.6766	71.1121	63.4653	67.7811	88.5023	82.5187	69.572	70.6478				
D12	67.1283	83.334	73.2689	75.5041	91.32	84.7817	75.3577	66.5002	62.495			
D15	78.0049	74.5194	70.5577	67.7507	94.6199	84.026	85.7638	77.6787	78.7367	79.6789		
D16	97.414	101.413	87.7375	94.0377	91.6336	92.627	90.5237	89.192	85.1862	88.1528	102.051	

Table B-15. An example of a distance matrix results from performing DTW algorithm on all signals 300 frames before left turns.

DRIVERS	D1	D2	D3	D4	D7	D8	D9	D10	D11	D12	D15	D16
D1												
D2	76.9869											
D3	80.4096	81.7646										
D4	66.5668	84.7649	80.9741									
D7	77.5829	87.9256	76.0633	80.9646								
D8	75.5646	85.6512	89.3406	84.3937	76.3555							
D9	82.5411	83.7344	79.8544	84.8809	76.717	79.9813						
D10	79.9043	88.7264	82.328	90.5699	89.2935	83.6095	97.1349					
D11	90.501	96.8265	95.2304	90.3329	100.343	89.5893	99.0561	96.2039				
D12	82.478	88.0825	71.297	90.2182	74.2948	83.1922	76.9821	88.4054	85.3761			
D15	80.2195	93.3252	81.2801	97.9038	81.0776	80.5134	91.8804	91.7265	94.9939	76.113		
D16	78.3812	86.1644	94.9889	87.2389	86.0389	72.4302	84.0195	97.4736	83.5687	92.882	93.8019	

Appendix C

Sample Implementation Code

Statistical Feature Extraction:

```
clear all;

D=[];   % contains speed signal of one driver before right turns
drivers = {'1','2','3','4','7','8','9','10','11','12','15','16'} ;

prompt = 'Do you want to extract features of all turns, left turns or
right turns? A/L/R [A]: ';
str = input(prompt,'s');
if isempty(str)
    str = 'A';
end

if str=='A'
    trns=[3,7,8,9,11,13,14,2,4,5,6,10,12,15,17,18,19];
    trnTyp='All';
elseif str=='L'
    trns = [3,7,8,9,11,13,14];
    trnTyp='Left';
else
    trns = [2,4,5,6,10,12,15,17,18,19];
    trnTyp='Right';
end

prompt = 'How many frames? 150/300/450 ';
noFr = input(prompt);

dNo=1;
for d=1:size(drivers,2)
    s = strcat('Driver',drivers{d},'All.csv');
    M = csvread(s);  %all data:
speed,gas,brake,leftSteer,rightSteer,leftSignal,rightSignal

    s = strcat('turn_',drivers{d},'.csv');
    T = csvread(s);  %all turns

    for i=trns

        v=[];g=[];b=[];ls=[];rs=[];acc=[];
        currfr = T(i,1);    %current frame
        k = 0;              %number of sequenses with non zero speed
        l = 0;              %counter
        v1 = M(currfr,1);   %v1-v0/t for caculating accelerate
        while k ~= noFr
            if M(currfr,1) ~= 0
                v = [M(currfr,1),v];    %speed
                g = [M(currfr,2),g];    %gas
```

```
                b = [M(currfr,3),b];     %brake
                ls = [M(currfr,4),ls];  %left steering wheel
                rs = [M(currfr,5),rs];  %right steering wheel
                sw = rs-ls;             %steering wheel
                if l == 30
                    l=0;
                    v0 = M(currfr,1);
                    acc = [v1-v0,acc];  %accelerate (every 30 frames)
                    v1=v0;
                end
                k=k+1;
                l=l+1;
            end
            currfr = currfr - 1;
        end

        v=v'; g=g'; b=b'; ls=ls'; rs=rs'; acc=acc';

        s=strcat('turn',num2str(i),'extracted',num2str(noFr),'.csv');
        if exist(s,'file')
            X = csvread(s);
        else
            X=[];
        end

        D=[X;
            mean(v),std(v),kurtosis(v),skewness(v),...
            mean(g),std(g),kurtosis(g),skewness(g),...
            mean(b),std(b),kurtosis(b),skewness(b),...
            mean(sw),std(sw),kurtosis(sw),skewness(sw),...
            mean(acc),std(acc),kurtosis(acc),skewness(acc)];

        csvwrite(s,D);
    end
    dNo=dNo+1;
end


total=zeros(12,20);
for i=trns
    s=strcat('turn',num2str(i),'extracted',num2str(noFr),'.csv');
    x=csvread(s);
    x(isnan(x)) = 0;
    total = total + x;
end

total=total/size(trns,2);
s=strcat(trnTyp,'TurnsExtracted',num2str(noFr),'.csv');
csvwrite(s,total);
```

Statistical feature extraction clustering:

```
clear all;

prompt = 'All turns, left turns or right turns? A/L/R [A]: ';
str = input(prompt,'s');
if isempty(str)
    str = 'A';
end

if str == 'A'
    trnTyp = 'All';
elseif str == 'L'
    trnTyp = 'Left';
else
    trnTyp = 'Right';
end

prompt = 'How many frames? 150/300/450 ';
noFr = input(prompt);

s=strcat(trnTyp,'TurnsExtracted',num2str(noFr),'.csv');
turns=csvread(s);


% Normalizing the feature values between 0 and 1
for i = 1:size(turns,2)
    turns(:,i)=(turns(:,i)-min(turns(:,i)))/(max(turns(:,i))-
min(turns(:,i))+0.0001);
end

turns(isnan(turns)) = 0;

% performing HCA using Ward linkage criteria and Euclidean Distance
Z = linkage(turns,'ward');

figure;
H = dendogram(Z, 'labels',{'Driver1  (37/M)','Driver2  (37/M)','Driver3
(41/F)','Driver4  (41/M)','Driver7  (31/F)','Driver8  (21/M)','Driver9
(21/F)','Driver10  (20/M)','Driver11  (22/F)','Driver12
(24/F)','Driver15  (44/F)','Driver16  (25/M)'});
set(H,'LineWidth',1); ax = gca; ax.XTickLabelRotation = 45;
```

References

Aarts, Letty, and Ingrid van Schagen. 2006. “Driving Speed and the Risk of Road Crashes: A Review.” *Accident Analysis & Prevention* 38 (2): 215–24. doi:10.1016/j.aap.2005.07.004.

Beauchemin, S. S., M. A. Bauer, T. Kowsari, and J. Cho. 2012. “Portable and Scalable Vision-Based Vehicular Instrumentation for the Analysis of Driver Intentionality.” *IEEE Transactions on Instrumentation and Measurement* 61 (2): 391–401. doi:10.1109/TIM.2011.2164854.

Berndt, Donald J., and James Clifford. 1994. “Using Dynamic Time Warping to Find Patterns in Time Series.” In *KDD Workshop*, 10:359–370. Seattle, WA. http://www.aaai.org/Library/Workshops/1994/ws94-03-031.php.

Carmona, Juan, Fernando García, David Martín, Arturo de la Escalera, and José María Armingol. 2015. “Data Fusion for Driver Behaviour Analysis.” *Sensors* 15 (10): 25968–25991.

Chandler, Robert E., Robert Herman, and Elliott W. Montroll. 1958. “Traffic Dynamics: Studies in Car Following.” *Operations Research* 6 (2): 165–184.

Chen, Shi-Huang, Jeng-Shyang Pan, and Kaixuan Lu. 2015. “Driving Behaviour Analysis Based on Vehicle OBD Information and AdaBoost Algorithms.” In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 1:18–20. http://www.iaeng.org/publication/IMECS2015/IMECS2015_pp102-106.pdf.

Choi, SangJo, JeongHee Kim, DongGu Kwak, Pongtep Angkititrakul, and John HL Hansen. 2007. “Analysis and Classification of Driver Behaviour Using in-Vehicle Can-Bus Information.” In *Biennial Workshop on DSP for In-Vehicle and Mobile Systems*, 17–19. https://www.researchgate.net/profile/Pongtep_Angkititrakul/publication/228936619_Analysis_and_Classification_of_Driver_Behaviour_using_In-Vehicle_CAN-Bus_Information/links/00b7d51c0e8f32c704000000.pdf.

Enev, Miro, Alex Takakuwa, Karl Koscher, and Tadayoshi Kohno. 2016. “Automobile Driver Fingerprinting.” *Proceedings on Privacy Enhancing Technologies* 2016 (1): 34–50.

Eren, H., S. Makinist, E. Akin, and A. Yilmaz. 2012. “Estimating Driving Behaviour by a Smartphone.” In *2012 IEEE Intelligent Vehicles Symposium (IV)*, 234–39. doi:10.1109/IVS.2012.6232298.

Fildes, B. N., G. Rumbold, and A. Leening. 1991. “Speed Behaviour and Drivers’ Attitude to Speeding.” *Monash University Accident Research Centre, Report* 16. https://www.researchgate.net/profile/Brian_Fildes/publication/238493381_Speed_Behaviour_and_Drivers'_Attitudes_to_Speeding/links/546ac8fe0cf20dedafd38e8d.pdf.

Higgs, B., and M. Abbas. 2013. “A Two-Step Segmentation Algorithm for Behavioral Clustering of Naturalistic Driving Styles.” In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 857–62. doi:10.1109/ITSC.2013.6728339.

———. 2015. “Segmentation and Clustering of Car-Following Behavior: Recognition of Driving Patterns.” *IEEE Transactions on Intelligent Transportation Systems* 16 (1): 81–90. doi:10.1109/TITS.2014.2326082.

Johnson, D. A., and M. M. Trivedi. 2011. “Driving Style Recognition Using a Smartphone as a Sensor Platform.” In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 1609–15. doi:10.1109/ITSC.2011.6083078.

Kalsoom, R., and Z. Halim. 2013. “Clustering the Driving Features Based on Data Streams.” In *Multi Topic Conference (INMIC), 2013 16th International*, 89–94. doi:10.1109/INMIC.2013.6731330.

Meinard Müller. 2007. “Dynamic Time Warping.” In *Information Retrieval for Music and Motion*, 69–84. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-74048-3_4.

Miyajima, Chiyomi, Yoshihiro Nishiwaki, Koji Ozawa, Toshihiro Wakita, Katsunobu Itou, Kazuya Takeda, and Fumitada Itakura. 2007. “Driver Modeling Based on Driving Behavior and Its Evaluation in Driver Identification.” *Proceedings of the IEEE* 95 (2): 427–437.

Ohta, Hiroo. 1993. “Individual Differences in Driving Distance Headway.” *Vision in Vehicles* 4: 91–100.

Sakoe, H., and S. Chiba. 1978. “Dynamic Programming Algorithm Optimization for Spoken Word Recognition.” *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26 (1): 43–49. doi:10.1109/TASSP.1978.1163055.

Wang, Quan, 2015. Dynamic Time Warping (DTW) (https://www.mathworks.com/matlabcentral/fileexchange/43156-dynamic-time-warping--dtw-), MATLAB Central File Exchange. Retrieved May 17, 2016.

Wu, Chaozhong, Chuan Sun, Duanfeng Chu, Zhen Huang, Jie Ma, and Haoran Li. 2016. “Clustering of Several Typical Behavioral Characteristics of Commercial Vehicle Drivers Based on GPS Data Mining.” *Transportation Research Record: Journal of the Transportation Research Board* 2581 (January): 154–63. doi:10.3141/2581-18.

Curriculum Vitae

Name:	Mahboubeh Zardosht
Post-secondary Education and Degrees:	Shiraz University Shiraz, Iran 2005-2009 B.Sc. University of Isfahan Isfahan, Iran 2009-20012 M.Sc. The University of Western Ontario London, Ontario, Canada 2014-2016 M.Sc.
Honors and Awards:	Western Graduate Research Scholarship 2014-2015
Related Work Experience	Research Assistant The University of Western Ontario 2014-2016 Teaching Assistant The University of Western Ontario 2014-2016 Lecturer Javid Institute of Higher Education 2014-2014 Lecturer Institute of Applied Science and Technology 2012-2013