NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF CIVIL ENGINEERING
DEPARTMENT OF TRANSPORTATION PLANNING & ENGINNERING

DOCTORAL DISSERTATION

DEEP REINFORCEMENT LEARNING TRAFFIC MODELS FOR PERSONALIZED DRIVING RECOMMENDATIONS

Eleni G. Mantouka
Civil & Transportation Engineer, NTUA

SUPERVISOR: E. I. VLAHOGIANNI, ASSOCIATE PROFESSOR
CO-SUPERVISOR: I. GOLIAS, PROFESSOR
CO-SUPERVISOR: G. YANNIS, PROFESSOR

Athens, November 2021

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research – 2nd Cycle" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

Contact Details

To anyone who is willing to learn more about this PhD dissertation, please feel free to contact me using the following contact details:

Eleni G. Mantouka

National Technical University of Athens (NTUA), Department of Transportation Planning and Engineering, 5 Heroon Polytechniou Str., Zografou Campus, GR-15773, Athens, Greece

Professional e-mail: elmant@central.ntua.gr

Personal e-mail: elmantouka@gmail.com

Phone (NTUA): (+30) 2107721367

Ευχαριστίες

Ολοκληρώνοντας τη διδακτορική μου διατριβή οφείλω πρωτίστως να ευχαριστήσω την επιβλέπουσά μου κα Ελένη Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια ΕΜΠ, για την ευκαιρία που μου έδωσε να ξεκινήσω το ταξίδι μου στην έρευνα με ένα τόσο ενδιαφέρον θέμα. Ευχαριστώ για την εμπιστοσύνη που μου δείξατε, την καθοδήγηση, την εξαιρετική συνεργασία και για τα μαθήματα ζωής που μου διδάσκετε.

Ένα μεγάλο ευχαριστώ θέλω να πω στον κύριο Γιώργο Γιαννή, Καθηγητή ΕΜΠ για την αμέριστη υποστήριξή του σε όλη τη διάρκεια εκπόνησης της διδακτορικής διατριβής και για τα εποικοδομητικά του σχόλια, και εξίσου να ευχαριστήσω τον κύριο Ιωάννη Γκόλια, Καθηγητή ΕΜΠ για τις επιστημονικές συζητήσεις που κάναμε επί του θέματος της διατριβής.

Επιπλέον, θα ήθελα να ευχαριστήσω τα υπόλοιπα μέλη της επταμελούς επιτροπής κ.κ. Κωνσταντίνο Κεπαπτσόγλου, Νικόλα Γερολιμίνη, Ιωάννη Παπαμιχαήλ και Κωνσταντίνα Γκριτζά, για τις εύστοχες παρατηρήσεις τους που συνέβαλαν καθοριστικά στην τελική μορφή της διατριβής.

Θέλω επίσης να ευχαριστήσω την OSeven Telematics για την παροχή των δεδομένων, και προσωπικά τον κύριο Πέτρο Φορτσάκη για την άψογη συνεργασία μας.

Το ταξίδι αυτό της έρευνας δε θα ήταν ίδιο χωρίς τη φοβερή ομάδα του 201 που έκανε τις μέρες μου στο γραφείο ξεχωριστές, με πολλά γέλια και στιγμές που θα μου μείνουν αξέχαστες. Ευχαριστώ όλους τους συναδέλφους και φίλους μου, και ειδικά Φωτεινή, Αποστόλη, Άκη, Δημήτρη, Βασίλη, Μάνο και Μάνο, Αρετή και Κατερίνα, σας ευχαριστώ πολύ που ομορφύνατε αυτό το ταξίδι!

Ένα ξεχωριστό ευχαριστώ οφείλω να πω στο φίλο και συνάδελφό μου Χάρη για την πολύτιμη βοήθειά του στις πολλές προσομοιώσεις που κάναμε στο πλαίσιο της διατριβής!

Το μεγαλύτερο ευχαριστώ από όλα το οφείλω στην οικογένειά μου, τους γονείς μου Γιώργο και Γεωργία, την αδερφή μου Βάσω που στάθηκαν δίπλα μου σε κάθε βήμα, πίστεψαν σε μένα και μου προσέφεραν απλόχερα την υποστήριξη και την αγάπη τους. Σας ευχαριστώ πολύ, δε θα τα κατάφερνα χωρίς εσάς.

Κλείνοντας, θέλω να ευχαριστήσω τον άντρα της ζωής μου. Παναγιώτη σ' ευχαριστώ που ομορφαίνεις την καθημερινότητά μου, με παροτρύνεις να γίνομαι καλύτερη, πιστεύεις σε μένα και ενθαρρύνεις τα όνειρά μου.

Abstract

The aim of this doctoral dissertation is initially to develop an inclusive personalized driving recommendation framework which first identifies each individual's driving style and then, proposes customized driving actions that mitigate aggressiveness and riskiness, and, thereafter, to assess the impact of applying such a personalized decision support system through microsimulation by properly adjusting traffic models. The methodological approach is based on a mixture of unsupervised learning and Deep Reinforcement Learning algorithms while exploiting an always increasing naturalistic driving dataset that emerged from crowd-sensing. Using a variety of variables which describe driving behavior each individual's driving behavior is identified both on a trip- and user-level. Findings revealed that there are 6 driving profiles that describe the overall behavior a driver has during their trip, which spans from safe to aggressiveness and distraction during driving. Subsequently, the average behavior of each driver was estimated and further exploited to separate drivers into groups. Then, two Reinforcement Learning agents were trained, each one corresponding to a specific group of drivers with common behavior, to determine the optimal behavior alteration for each driver given the way they have drove over their last trip. Specifically, the two agents correspond to cautious and more aggressive drivers respectively. The results of applying the controllers into real world data have shown that, although given the same driving characteristics of a trip, indicating the same driving behavior, the two controllers provide different driving suggestions, both of them lead to safer driving actions. Finally, a microsimulation scenario was set in order to assess traffic conditions, emissions and safety before and after applying the recommendation system. Findings have revealed that when each individual driver improves their behavior accordingly to the system's recommendation, although traffic conditions are not improved, the calculation of key safety and environmental performance indicators unveiled a significant reduction of both conflicts and harmful emissions. Results could be exploited within the framework of an advanced active cruise control system, in the development of enhanced behavioral models or could even lead to the revision of policy measures that utilize driving behavior as a key controller of traffic management.

Keywords: *driving behavior, reinforcement learning, k-means clustering, driving recommendations, personalization, microsimulation, driving safety, big data*

Σύνοψη

Ο στόχος της παρούσας διδακτορικής διατριβής είναι, αρχικά, να αναπτύξει μια ολοκληρωμένη προσέγγιση για τη δημιουργία ενός εξατομικευμένου συστήματος παροχής συστάσεων οδήγησης το οποίο πρώτα θα ανιχνεύει τη συμπεριφορά οδήγησης και στη συνέχεια θα προτείνει εξατομικευμένες ενέργειες που μετριάζουν την επιθετικότητα και το ρίσκο κατά την οδήγηση, και έπειτα, να εκτιμήσει τις επιπτώσεις της εφαρμογής ενός τέτοιου συστήματος μέσω προσομοίωσης προσαρμόζοντας κατάλληλα τις παραμέτρους του κυκλοφοριακού μοντέλου. Η μεθοδολογική προσέγγιση βασίζεται σε ένα μείγμα αλγορίθμων μη επιβλεπόμενης και ενισχυτικής μάθησης, στα πλαίσια των οποίων αξιοποιείται μια βάση δεδομένων πραγματικής οδήγησης που συλλέγονται από το πλήθος. Χρησιμοποιώντας μια ποικιλία μεταβλητών που περιγράφουν τον τρόπο οδήγησης, προσδιορίζεται η συμπεριφορά οδήγησης τόσο σε επίπεδο διαδρομής όσο και σε επίπεδο χρήστη. Η ανάλυση ανέδειξε 6 προφίλ οδήγησης που περιγράφουν τη συνολική συμπεριφορά ενός οδηγού κατά τη διάρκεια της διαδρομής του, και εκτείνονται από την ασφαλή έως την επιθετική οδήγηση και την οδήγηση με απόσπαση προσοχής. Στη συνέχεια, η μέση συμπεριφορά κάθε οδηγού εκτιμήθηκε και αξιοποιήθηκε περαιτέρω για να διαχωριστούν οι οδηγοί σε ομάδες με κοινά χαρακτηριστικά. Δύο ξεχωριστοί πράκτορες Ενισχυτικής μάθησης εκπαιδεύτηκαν, ένας για κάθε ομάδα χρηστών, οι οποίοι είναι σε θέση να προσδιορίζουν τη βέλτιστη αλλαγή στη συμπεριφορά κάθε οδηγού δεδομένου του τρόπου που ο ίδιος οδήγησε στην προηγούμενη διαδρομή του. Πιο συγκεκριμένα, οι δύο πράκτορες αντιστοιχούν στους τυπικούς οδηγούς και στους πιο επιθετικούς/μη ασφαλείς οδηγούς, αντίστοιχα. Τα αποτελέσματα της αξιοποίησης των συστάσεων οδήγησης έδειξαν ότι, δεδομένων των ίδιων χαρακτηριστικών οδήγησης για μια διαδρομή, που υποδηλώνουν ίδια συμπεριφορά οδήγησης, οι δύο πράκτορες παράγουν διαφορετικές συστάσεις οι οποίες ταιριάζουν με τις προτιμήσεις της εκάστοτε ομάδας οδηγών, παρόλο που και οι δύο οδηγούν σε ασφαλέστερες ενέργειες οδήγησης. Τέλος, δύο κύκλοι προσομοίωσης πραγματοποιήθηκαν προκειμένου να αξιολογηθούν οι επιπτώσεις της εφαρμογής ενός τέτοιου συστήματος τόσο στην κυκλοφορία, όσο και στην οδική ασφάλεια και το περιβάλλον. Τα ευρήματα έδειξαν ότι εάν όλοι οι οδηγοί βελτιώσουν τη δική τους συμπεριφορά οδήγησης τότε, παρόλο που οι συνθήκες κυκλοφορίας δε βελτιώνονται, παρατηρούνται σημαντικές μειώσεις τόσο στις εκπομπές βλαβερών αερίων όσο και στις εμπλοκές ανάμεσα στα οχήματα. Τα αποτελέσματα αυτής της έρευνας μπορούν να αξιοποιηθούν στα πλαίσια ενός προηγμένου συστήματος υποβοήθησης του οδηγού, στην ανάπτυξη προχωρημένων προτύπων συμπεριφοράς ή ακόμη και να οδηγήσουν στην αναθεώρηση των πολιτικών που χρησιμοποιούν τη συμπεριφορά οδήγησης ως κινητήριο δύναμη για τη διαχείριση της κυκλοφορίας.

Λέξεις – Κλειδιά: *συμπεριφορά οδήγησης, ενισχυτική μάθηση, ομαδοποίηση k-means, συστάσεις οδήγησης, εξατομίκευση, προσομοίωση, οδηγική ασφάλεια, μεγάλα δεδομένα*

Table of contents

Table of figures

Table of tables

List of abbreviations

AV(s)	Autonomous Vehicle(s)
AI	Artificial Intelligence
CAV(s)	Connected and Autonomous Vehicle(s)
c-ITS	Cooperative Intelligent Transportation Systems
CO	Carbon Monoxide
CO_2	Carbon Dioxide
DDPG	Deep Deterministic Policy Gradient
GDPR	General Data Protection Regulation
GMM	Gaussian Mixture Model
GPS	Global Positioning System
IoT	Internet of Things
ITS	Intelligent Transportation Systems
KPI(s)	Key Performance Indicator(s)
MaaS	Mobility as a Service
MFD	Macroscopic Fundamental Diagram
NO_x	Oxides of Nitrogen
PAYD	Pay-As-You-Drive
PHYD	Pay-How-You-Drive
PM_x	Particulate Matter
PT	Public Transport
RL	Reinforcement Learning
SSAM	Surrogate Safety Assessment Model
SUMO	Simulation of Urban MObility
SVM	Support Vector Machines
UBI	Usage-based Insurance

Περίληψη Διδακτορικής Διατριβής

Εισαγωγή

Το αστικό δίκτυο μεταφορών έχει αλλάξει δραστικά τα τελευταία χρόνια λόγω των ποικίλων εναλλακτικών μετακίνησης που προσφέρονται στους μετακινούμενους, οι οποίες μαζί με τις καινοτομίες στον τομέα των Τεχνολογιών Πληροφορίας και Επικοινωνιών (ΤΠΕ) ανατρέπουν τις πολιτικές διαχείρισης της κυκλοφορίας και του δικτύου. Οι υπεύθυνοι χάραξης πολιτικής καλούνται να επανεξετάσουν τα ισχύοντα μέτρα και τις στρατηγικές διαχείρισης του δικτύου ώστε να λάβουν υπόψη τις απαιτήσεις για αυτοματοποίηση και συνεργατικότητα των σύγχρονων υπηρεσιών μετακίνησης, αλλά και τις συνεχώς αυξανόμενες ανάγκες για πράσινες και βιώσιμες λύσεις κινητικότητας. Ακόμα και σε αυτό το πολύπλοκο και διαρκώς μεταβαλλόμενο περιβάλλον, ο οδηγός βρίσκεται στο επίκεντρο του ενδιαφέροντος. Ως εκ τούτου, η κατανόηση της λήψης αποφάσεων κατά την οδήγηση, καθώς και η διερεύνηση των συνηθειών οδήγησης που υιοθετούν οι οδηγοί παραμένει ενεργό πεδίο έρευνας.

Η επικίνδυνη οδήγηση έχει συνδεθεί με αυξημένο ρίσκο και γι' αυτό η βελτίωση της συμπεριφοράς οδήγησης θεωρείται κρίσιμη για τη βελτίωση της οδικής ασφάλειας. Επιπλέον, προηγούμενες έρευνες έχουν υπονοήσει ότι η βελτίωση της ατομικής συμπεριφοράς οδήγησης μπορεί να επιφέρει και βελτίωση των συνθηκών κυκλοφορίας. Παρόλα αυτά όμως, δεν έχουν παρασχεθεί στοιχεία που να μπορούν να υποστηρίξουν αυτή την πεποίθηση και οι συνέπειες της προσαρμογής της ατομικής συμπεριφοράς οδήγησης σε επίπεδο δικτύου εξακολουθούν να παραμένουν ασαφείς.

Σε αυτό το πλαίσιο, τα δύο βασικά κίνητρα για την εκπόνηση της διδακτορικής διατριβής υπήρξαν τα ακόλουθα: 1) η ανάγκη σχεδιασμού ενός συστήματος υποβοήθησης του οδηγού που θα αντιμετωπίζει τον ίδιο ως μονάδα και θα προσαρμόζεται στις προσωπικές ανάγκες και προτιμήσεις του και 2) η ανάγκη εκτίμησης του πραγματικού αντικτύπου της εφαρμογής ενός συστήματος παροχής εξατομικευμένων συστάσεων οδήγησης στις διάφορες πτυχές του οδικού δικτύου.

Στόχοι της διατριβής και ερευνητικά ερωτήματα

Ο κύριος στόχος της παρούσας διατριβής είναι να σχεδιάσει ένα εξατομικευμένο σύστημα συστάσεων οδήγησης το οποίο βασίζεται σε αλγόριθμους βαθιάς ενισχυτικής μάθησης και στοχεύει στη βελτίωση της συμπεριφοράς οδήγησης μέσω του μετριασμού της επιθετικότητας και άλλων μη ασφαλών συνηθειών οδήγησης. Στη συνέχεια, αξιολογούνται οι επιπτώσεις του ελέγχου της ατομικής συμπεριφοράς οδήγησης όσον αφορά την απόδοση του δικτύου και την οδική ασφάλεια, καθώς και τα επίπεδα επιβλαβών εκπομπών, προσαρμόζοντας κατάλληλα τις παραμέτρους των κυκλοφοριακών μοντέλων σε ένα ευρύ σενάριο προσομοίωσης της κυκλοφορίας. Αυτός ο στόχος της διδακτορικής διατριβής μπορεί να αποδομηθεί σε τρεις επιμέρους στόχους:

1. Αξιοποίηση δεδομένων πραγματικής οδήγησης που συλλέγονται από αισθητήρες κινητών τηλεφώνων για την κατανόηση της συμπεριφοράς οδήγησης
2. Ανάπτυξη ενός πλαισίου παροχής συστάσεων για βελτιωμένη οδήγηση συμβατό με τις αρχές της θεωρίας της κυκλοφοριακής ροής
3. Αξιολόγηση της επιρροής του συστήματος στην κυκλοφορία, την ασφάλεια και τις εκπομπές.

Η έννοια της ανάλυσης της συμπεριφοράς οδήγησης δεν είναι καινούργια και για αυτό το λόγο διεξήχθη μια εκτενής ανασκόπηση της σχετικής βιβλιογραφίας με στόχο αφενός να εντοπιστούν τα κενά της έρευνας και αφετέρου να επισημανθούν οι προκλήσεις και οι περιορισμοί που προκύπτουν όταν αξιοποιούνται δεδομένα που προέρχονται από το πλήθος. Η ανασκόπηση της βιβλιογραφίας είχε ως αποτέλεσμα τη διαμόρφωση των ακόλουθων ερευνητικών ερωτημάτων:

1. Ποια είναι τα κύρια προφίλ οδήγησης που καλύπτουν το ευρύ φάσμα της συμπεριφοράς οδήγησης και πώς μπορούν να εντοπιστούν μέσω της αξιοποίησης δεδομένων που συλλέγονται μέσω έξυπνων κινητών τηλεφώνων;
2. Μπορεί η συνολική συμπεριφορά οδήγησης των οδηγών να κατηγοριοποιηθεί σε ομάδες που θα εμφανίζουν κοινά χαρακτηριστικά οδήγησης, και αν ναι, σε ποιο βαθμό μπορούν να κατηγοριοποιηθούν αυτές οι συμπεριφορές;
3. Θα μπορούσαν οι τεχνικές Τεχνητής Νοημοσύνης (Artificial Intelligence) να αξιοποιηθούν στα πλαίσια ενός συστήματος παραγωγής συστάσεων στους οδηγούς και να εξασφαλίσουν τον απαιτούμενο βαθμό εξατομίκευσης των ενεργειών οδήγησης που προτείνονται σε κάθε χρήστη;
4. Ποιος είναι ο πιο κατάλληλος αλγόριθμος Ενισχυτικής Μάθησης (Reinforcement Learning) που μπορεί να υποστηρίξει τη διαδικασία λήψης αποφάσεων του ατόμου;
5. Υπάρχει σύνδεση μεταξύ της ενίσχυσης της ευαισθητοποίησης του ατόμου και της καθολικής βελτίωσης του δικτύου; Σε ποιο βαθμό μπορεί η βελτίωση της συμπεριφοράς οδήγησης να επηρεάσει τις συνθήκες του δικτύου;
6. Τι είδους επιπτώσεις θα είχε η διαχείριση της ατομικής συμπεριφοράς στην οδήγηση και την ασφάλεια;
7. Πώς επηρεάζονται οι εκπομπές από τον έλεγχο της ατομικής συμπεριφοράς οδήγησης; Υπάρχει σημαντική αλλαγή στις περιβαλλοντικές συνθήκες όταν οι οδηγοί βελτιώνουν τη συμπεριφορά τους;

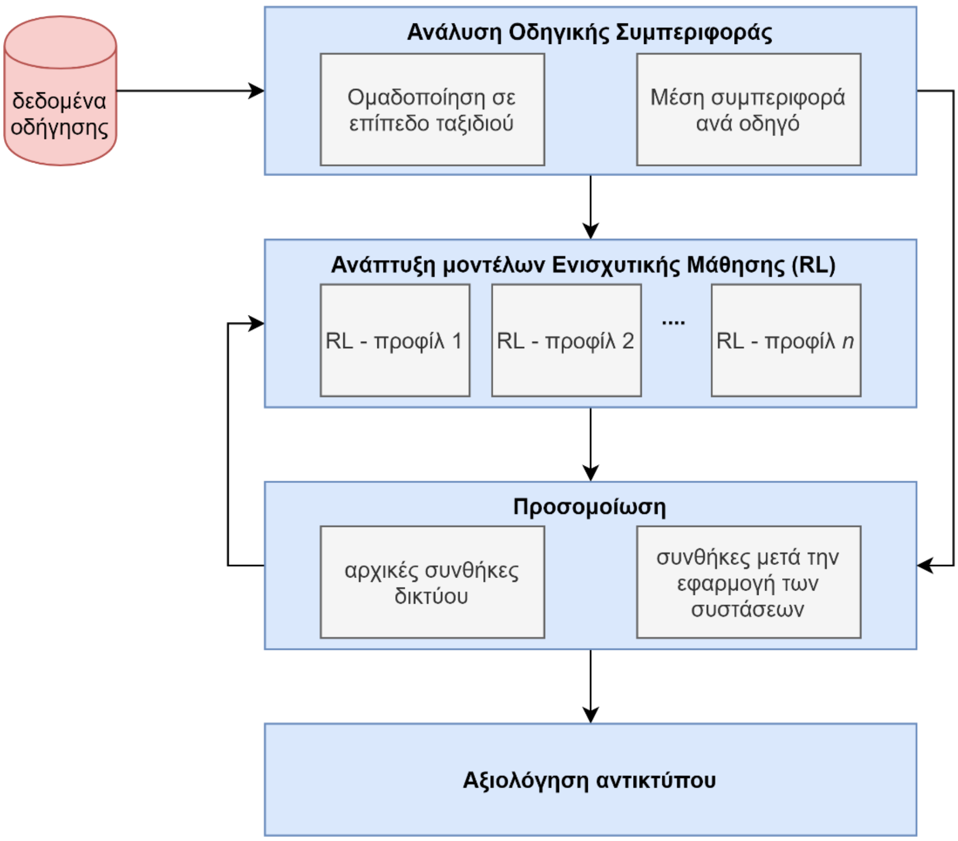
Μεθοδολογική προσέγγιση

Το μεθοδολογικό πλαίσιο για την παροχή συστάσεων οδήγησης, που προτείνεται στο πλαίσιο αυτής της διατριβής, είναι βασικά ένα σύστημα υποστήριξης αποφάσεων για τους οδηγούς που στοχεύει στο μετριασμό της επιθετικότητας και της ανάληψης ρίσκου. Η οδήγηση είναι μια πολύπλοκη εργασία, δεδομένου ότι απαιτεί από τον οδηγό να λάβει τόσο στρατηγικές όσο και δυναμικές αποφάσεις καθώς και να προσαρμόσει τη συμπεριφορά του στις εκάστοτε συνθήκες του δικτύου. Σε αντίθεση με τα ήδη αναπτυγμένα συστήματα υποβοήθησης του οδηγού (Advanced Driving Assistance Systems – ADAS), το προτεινόμενο σύστημα έχει τα ακόλουθα τρία καινοτόμα χαρακτηριστικά:

- Είναι εξατομικευμένο, που σημαίνει ότι προτείνει τις καλύτερες ενέργειες οδήγησης σε κάθε άτομο λαμβάνοντας υπόψη τις προσωπικές απαιτήσεις και προτιμήσεις οδήγησης.
- Έχει αυτο-γνωσία (self-awareness), που σημαίνει ότι το σύστημα λαμβάνει υπόψη την προηγούμενη συμπεριφορά κάθε μεμονωμένου οδηγού προκειμένου να προτείνει τις καταλληλότερες συστάσεις οδήγησης.
- Είναι αυτόνομο, που σημαίνει ότι δεν απαιτεί καμία εξωγενή πληροφορία από το δίκτυο ή την κυκλοφορία. Οι συστάσεις οδήγησης στοχεύουν στη βελτίωση της ατομικής συμπεριφοράς οδήγησης στον πυρήνα της, δηλαδή στις αποφάσεις επιτάχυνσης και επιβράδυνσης.

Η ανάπτυξη του συστήματος συστάσεων βασίζεται σε έναν αλγόριθμο Βαθιάς Ενισχυτικής Μάθησης ο οποίος είναι ικανός να παράγει τη βέλτιστη μεταβολή της συμπεριφοράς για κάθε οδηγό δεδομένου του τρόπου με τον οποίο οδήγησε στην τελευταία διαδρομή του.

Προκειμένου να απαντηθούν τα ερευνητικά ερωτήματα που τέθηκαν και να επιτευχθεί ο πρωταρχικός στόχος της διατριβής, προτείνεται ένα περιεκτικό μεθοδολογικό πλαίσιο το οποίο βασίζεται σε ένα μείγμα αλγορίθμων μάθησης χωρίς επίβλεψη και βαθιάς ενισχυτικής μάθησης όπως απεικονίζεται στο Σχήμα I.



Σχήμα I. Μεθοδολογικό πλαίσιο

Ξεκινώντας από πρωτογενείς μετρήσεις θέσης, επιτάχυνσης και ταχύτητας, όπως αυτές συλλέχθηκαν μέσω των αισθητήρων των έξυπνων κινητών τηλεφώνων χρησιμοποιώντας την εφαρμογή τηλεματικής της Oseven (www.oseven.io) προσδιορίστηκαν παράμετροι που περιγράφουν τη συμπεριφορά κατά την οδήγηση, τόσο βραχυπρόθεσμα, όσο και μακροπρόθεσμα. Στη συνέχεια, αυτές οι παράμετροι χρησιμοποιούνται σε ένα πλαίσιο μάθησης χωρίς επίβλεψη για τον εντοπισμό προφίλ οδήγησης που μπορούν να περιγράψουν τη συνολική συμπεριφορά οδήγησης κάθε οδηγού (E1, E2). Η συμπεριφορά οδήγησης ορίζεται σε δύο επίπεδα:

- σε επίπεδο διαδρομής, το οποίο αντιστοιχεί στον τρόπο με τον οποίο ο οδηγός οδήγησε σε μια συγκεκριμένη διαδρομή, και
- σε επίπεδο χρήστη, το οποίο αντιστοιχεί στη συνολική συμπεριφορά οδήγησης ενός συγκεκριμένου οδηγού σε όλα τα ταξίδια του (αποτύπωμα οδήγησης).

Για τον προσδιορισμό των διάφορων προφίλ οδήγησης που διέπουν τη συμπεριφορά οδήγησης σε κάθε διαδρομή, εφαρμόζεται ένας αλγόριθμος ομαδοποίησης k-means σε δύο διακριτά επίπεδα. Στο πρώτο επίπεδο εντοπίζεται η επιθετικότητα κατά την οδήγηση, ενώ στο δεύτερο επίπεδο προσδιορίζονται επιπρόσθετα άλλες μη ασφαλείς συμπεριφορές οδήγησης, όπως η οδήγηση πάνω από το όριο ταχύτητας (ανάληψη ρίσκου) και η απόσπαση προσοχής. Αποτέλεσμα αυτής της διαδικασίας είναι η αντιστοίχιση κάθε εκτελεσθείσας διαδρομής με ένα συγκεκριμένο προφίλ οδήγησης (E1). Στη συνέχεια, υπολογίζοντας τα στατιστικά χαρακτηριστικά όλων των διαδρομών κάθε οδηγού, προσδιορίζεται η μέση συμπεριφορά του καθενός (E2).

Από τη στιγμή που θα οριστεί η μέση συμπεριφορά των οδηγών, οι τελευταίοι χωρίζονται σε ομάδες με τέτοιο τρόπο, ώστε σε κάθε ομάδα να ανήκουν οδηγοί με κοινά χαρακτηριστικά οδήγησης. Έπειτα, σχεδιάζεται το σύστημα παροχής συστάσεων οδήγησης το οποίο ενσωματώνει έναν αλγόριθμο Ενισχυτικής Μάθησης που έχει ως στόχο να «μάθει» τη βέλτιστη πολιτική και να προτείνει την κατάλληλη δράση που οδηγεί στην καλύτερη δυνατή συμπεριφορά οδήγησης (E3). Στο συγκεκριμένο πρόβλημα, η δράση που προτείνεται αφορά στην προσαρμογή των κινηματικών χαρακτηριστικών του οχήματος δηλαδή της ταχύτητας και της επιτάχυνσης τα οποία εκτείνονται σε ένα συνεχές εύρος τιμών. Για αυτό το λόγο, η επιλογή του κατάλληλου αλγορίθμου ενισχυτικής μάθησης θα πρέπει να ικανοποιεί την ανάγκη διαχείρισης δράσεων που λαμβάνουν συνεχείς τιμές (E4). Έτσι, επιλέγεται η ανάπτυξη ενός μοντέλου που ακολουθεί την προσέγγιση "actor-critic" και συγκεκριμένα αναπτύσσεται ο αλγόριθμος Deep Deterministic Policy Gradient, ο οποίος ενσωματώνει δύο βαθιά νευρωνικά δίκτυα, οι υπερπαράμετροι και η δομή των οποίων προκύπτουν μετά από συγκριτική αξιολόγηση των πιθανών συνδυασμών. Ο αλγόριθμος εκπαιδεύεται χρησιμοποιώντας ακολουθίες διαδρομών οδήγησης του ίδιου οδηγού ως είσοδο, ενώ η έξοδος του αλγορίθμου, δηλαδή η προτεινόμενη ενέργεια, είναι η βέλτιστη αλλαγή στην επιτάχυνση κάθε οδηγού, δεδομένου του τρόπου με τον οποίο οδήγησε στην προηγούμενη διαδρομή του.

Η δομή του συστήματος είναι τέτοια ώστε να υπάρχει πλήρης αντιστοίχιση με λογικές μικροσκοπικής προτυποποίησης και ελέγχου της κυκλοφοριακής ροής. Σε αντιστοίχιση με τα ευρέως διαδεδομένα πρότυπα ακολουθούντος οχήματος (car following models), ο προτεινόμενος αλγόριθμος λειτουργεί ως μια συνάρτηση εκτίμησης και πρόβλεψης της επιτάχυνσης με την οποία το όχημα πρέπει να κινηθεί.

Το τελευταίο στάδιο της μεθοδολογίας που ακολουθήθηκε περιλαμβάνει την αξιολόγηση της επιρροής που θα έχει η βελτίωση της ατομικής συμπεριφοράς τόσο στο σύνολο της κυκλοφορίας, όσο και στην οδική ασφάλεια και το περιβάλλον (E5, E6, E7). Για το σκοπό αυτό ορίστηκε ένα σενάριο προσομοίωσης που αφορά στην πρωινή ώρα αιχμής αξιοποιώντας το οδικό δίκτυο της Αθήνας. Για την εκτίμηση της επιρροής ακολουθήθηκε μια προσέγγιση «πριν και μετά», όπου στις αρχικές συνθήκες του δικτύου η κυκλοφορία απαρτίζεται από οχήματα που κινούνται με βάση τα χαρακτηριστικά των προφίλ οδήγησης που ορίστηκαν στο πρώτο βήμα της μεθοδολογίας ενώ στο δεύτερο γύρο προσομοίωσης τα οχήματα κινούνται με βάση τις συστάσεις που παρήγαγε το μοντέλο ενισχυτικής μάθησης για κάθε οδηγό.

Δεδομένα οδήγησης

Για τις ανάγκες της εν λόγω έρευνας αξιοποιήθηκαν δεδομένα πραγματικής οδήγησης που συλλέχθηκαν μέσω μιας εφαρμογής τηλεματικής που αναπτύσσεται από την Oseven telematics. Η βάση δεδομένων περιλάμβανε 153.953 διαδρομές που πραγματοποιήθηκαν από 696 μοναδικούς οδηγούς από το Δεκέμβριο 2017 έως τον Αύγουστο 2019. Οι διαδρομές πραγματοποιήθηκαν στο οδικό δίκτυο της Ελλάδας, όμως η πλειοψηφία αυτών αφορά σε διαδρομές εντός του νομού Αττικής. Για κάθε διαδρομή, ήταν διαθέσιμες μια πληθώρα παραμέτρων που περιλαμβάνουν στατιστικά της επιτάχυνσης και της επιβράδυνσης, μετρήσεις ταχύτητας και δείκτες απόσπασης της προσοχής όπως η διάρκεια χρήσης του κινητού τηλεφώνου κατά την οδήγηση. Οι παράμετροι που χρησιμοποιήθηκαν στη συγκεκριμένη διατριβή παρουσιάζονται στον Πίνακα Ι.

Πίνακας Ι. Παράμετροι οδήγησης ανά διαδρομή

Όνομα μεταβλητής	Περιγραφή	Μονάδα μέτρησης
harsh_acc_per_min	Μέσος αριθμός απότομων επιταχύνσεων ανά λεπτό	συμβάντα/λεπτό
acc_avg	Μέση επιτάχυνση	m/s^2
acc_std	Τυπική απόκλιση επιτάχυνσης	m/s^2
acc_q90	90% της επιτάχυνσης	m/s^2
acc_max	Μέγιστη επιτάχυνση	m/s^2
harsh_brk_per_min	Μέσος αριθμός απότομων επιβραδύνσεων ανά λεπτό	συμβάντα/λεπτό
dec_avg	Μέση επιβράδυνση	m/s^2
dec_std	Τυπική απόκλιση επιβράδυνσης	m/s^2
dec_q90	90% της επιβράδυνσης	m/s^2
dec_max	Μέγιστη επιβράδυνση	m/s^2
speed_max	Μέγιστη ταχύτητα	km/h
mbu	Ποσοστό χρόνου οδήγησης με χρήση κινητού	%
speeding_percentage	Ποσοστό χρόνου οδήγησης με ταχύτητα πάνω από το όριο	%

Όλα τα δεδομένα που αξιοποιήθηκαν, δόθηκαν από την Oseven σε πλήρως ανώνυμη μορφή. Ο Πίνακας II παρουσιάζει τα βασικά χαρακτηριστικά του δείγματος που χρησιμοποιήθηκε.

Πίνακας II. Βασικά χαρακτηριστικά του δείγματος της διατριβής

	Σύνολο	Ασφαλείς διαδρομές	Μη ασφαλείς διαδρομές
Αριθμός διαδρομών	153.953	66.566	87.387
Αριθμός οδηγών	696	197	499
Μέσος αριθμός διαδρομών ανά οδηγό		221	
Ελάχιστος αριθμός διαδρομών ανά οδηγό		16	
Μέσος όρος χιλιομέτρων οδήγησης ανά οδηγό		2.510 km	

Ανάλυση συμπεριφοράς οδήγησης

Για την επίτευξη του πρώτου στόχου της διδακτορικής διατριβής, ο οποίος είναι η αναγνώριση της συμπεριφοράς οδήγησης από δεδομένα που προέρχονται από τους αισθητήρες των έξυπνων κινητών τηλεφώνων και συλλέγονται απευθείας από το πλήθος, εφαρμόστηκε ο αλγόριθμος ομαδοποίησης k-means σε δύο διακριτά επίπεδα.

Στο πρώτο επίπεδο της ομαδοποίησης, ο αριθμός των συστάδων ορίστηκε k=2 κι ο πίνακας αποστάσεων εκτιμήθηκε με βάση την Ευκλείδεια απόσταση. Οι παράμετροι που χρησιμοποιήθηκαν σε αυτό το επίπεδο της ομαδοποίησης περιγράφουν τον αριθμό των απότομων επιταχύνσεων και επιβραδύνσεων, καθώς και τα στατιστικά χαρακτηριστικά της επιτάχυνσης και της επιβράδυνσης. Τα σχετικά αποτελέσματα δίνονται στον Πίνακα III.

Πίνακας III. Αποτελέσματα πρώτου επιπέδου ομαδοποίησης

	Απότομες επιταχύνσεις ανά λεπτό	Απότομες επιβραδύνσεις ανά λεπτό	Μέση επιτάχυνση	Τυπική απόκλιση επιτάχυνσης	Μέγιστη επιτάχυνση	Μέση επιβράδυνση	Τυπική απόκλιση επιβράδυνσης	Μέγιστη επιβράδυνση	Αριθμός διαδρομών
Επιθετικές διαδρομές	0,150	0,2081	1,748	1,525	3,847	-1,968	1,843	-4,547	71.263
Μη επιθετικές διαδρομές	0,028	0,051	1,137	1,052	2,503	-1,282	1,286	-2,926	82.690

Σύμφωνα με τα κέντρα των ομάδων, οι διαδρομές μπορούν να διακριθούν σε επιθετικές και μη-επιθετικές, μιας και οι διαδρομές που ανήκουν στην πρώτη ομάδα χαρακτηρίζονται από επιθετικότητα κατά την οδήγηση όπως προκύπτει από τις μεγάλες επιταχύνσεις και επιβραδύνσεις, αλλά και το μεγαλύτερο αριθμό απότομων συμβάντων συγκριτικά με τις διαδρομές της δεύτερης ομάδας.

Το δεύτερο επίπεδο ομαδοποίησης k-means εφαρμόστηκε ξεχωριστά στις δύο ομάδες που προέκυψαν από το πρώτο επίπεδο ομαδοποίησης χρησιμοποιώντας δύο παραμέτρους οδήγησης: το ποσοστό οδήγησης με χρήση κινητού και το ποσοστό οδήγησης με ταχύτητα πάνω από το όριο ταχύτητας. Τα αποτελέσματα αυτού του δεύτερου επιπέδου ομαδοποίησης παρουσιάζονται στον Πίνακα IV.

Πίνακας IV. Αποτελέσματα δεύτερου επιπέδου ομαδοποίησης

	Ποσοστό οδήγησης με χρήση κινητού τηλεφώνου	Ποσοστό οδήγησης με ταχύτητα πάνω από το όριο	Αριθμός διαδρομών
Επιθετικές διαδρομές			
Απόσπαση προσοχής	0,511	0,062	4.505 (2,9%)
Επιθετικότητα	0,019	0,032	54.394 (35,3%)
Ανάληψη ρίσκου	0,023	0,269	12.364 (8%)
Μη-επιθετικές διαδρομές			
Ανάληψη ρίσκου	0,021	0,306	12.494 (8,1%)
Τυπική οδήγηση	0,014	0,029	66.566 (43,2%)
Απόσπαση προσοχής	0,514	0,057	3.630 (2,4%)

Στο δεύτερο στάδιο της ομαδοποίησης προκύπτουν πιο εμπλουτισμένα προφίλ οδήγησης. Συγκεκριμένα, αναγνωρίζεται η απόσπαση προσοχής κατά την οδήγηση από τις αυξημένες τιμές του ποσοστού χρήσης του κινητού τηλεφώνου και η ανάληψη ρίσκου από το μεγάλο ποσοστό χρόνου οδήγησης με ταχύτητα πάνω από το επιτρεπόμενο όριο. Οι διαδρομές που δε χαρακτηρίζονται από καμία από αυτές τις μη ασφαλείς συμπεριφορές αναφέρονται ως «επιθετικές» στην περίπτωση που στο πρώτο στάδιο χαρακτηρίστηκαν από επιθετικότητα και «τυπικές» στην περίπτωση που είχαν αρχικώς χαρακτηριστεί ως μη-επιθετικές διαδρομές.

Έτσι προκύπτουν έξι προφίλ οδήγησης που διέπουν τον τρόπο που μπορεί να οδηγήσει ένας οδηγός σε κάποια συγκεκριμένη διαδρομή του. Προκειμένου να διερευνηθεί περαιτέρω η συνολική συμπεριφορά των οδηγών, υπολογίζεται η μέση συμπεριφορά κάθε χρήστη εφαρμόζοντας έναν απλό κανόνα. Για κάθε μεμονωμένο οδηγό υπολογίζεται ο μέσος όρος των διαδρομών του, εφόσον η κάθε διαδρομή έχει συνδεθεί με ένα από τα έξι προφίλ οδήγησης αφού πρώτα τα τελευταία ταξινομήθηκαν με βάση την ασφάλεια κατά την οδήγηση, όπως φαίνεται στο Σχήμα II.

Σχήμα II. Κατηγοριοποίηση προφίλ οδήγησης για τον υπολογισμό της μέσης συμπεριφοράς οδήγησης

Πιο συγκεκριμένα, τα τέσσερα προφίλ οδήγησης που χαρακτηρίζονται από μη ασφαλή συμπεριφορά οδήγησης τοποθετήθηκαν στην κατηγορία 3, ενώ διαδρομές με επιθετική συμπεριφορά οδήγησης τοποθετήθηκαν στην κατηγορία 2 και τέλος, όσες διαδρομές ανήκουν στο τυπικό προφίλ οδήγησης αποτελούν την κατηγορία 1. Για κάθε μεμονωμένο οδηγό, υπολογίζεται ένας μέσος όρος των προφίλ όλων των διαδρομών του, και στην περίπτωση που ο μέσος όρος είναι μικρότερος από 1,5 συνεπάγεται μέτριο/τυπικό οδηγό όταν είναι μεγαλύτερος από 1,5 αναφέρεται σε απερίσκεπτους οδηγούς:

- μέσοι/τυπικοί οδηγοί: *μέσος όρος διαδρομών* $\leq$ *1,5*
- απερίσκεπτοι οδηγοί: *μέσος όρος διαδρομών* $>$ *1,5*

Εφαρμόζοντας αυτό τον κανόνα οι οδηγοί χωρίζονται σε δύο ομάδες, για κάθε μία από τις οποίες θα αναπτυχθεί διαφορετική έκδοση του αλγορίθμου παροχής συστάσεων οδήγησης με τρόπο ώστε ο αλγόριθμος να προσαρμόζεται στις προτιμήσεις οδήγησης του κάθε χρήστη. Με βάση τη στατιστική ανάλυση που πραγματοποιήθηκε, μέσος όρος ταξιδιών μικρότερος από 1,5 δείχνει ότι τουλάχιστον το 60% των ταξιδιών που εκτελεί ένας οδηγός χαρακτηρίζεται από "μέτρια" συμπεριφορά οδήγησης. Η προτεινόμενη μεθοδολογία προσδιορισμού του αποτυπώματος οδήγησης είναι ιδιαίτερα αυστηρή ως προς τον χαρακτηρισμό ενός οδηγού ως «τυπικού – ασφαλή» οδηγού προκειμένου να αποφευχθεί η πρόταση αλλαγών στη συμπεριφορά που ο ίδιος ο οδηγός είναι αδύνατο να ακολουθήσει καθώς θα είναι μακριά από τη δική του μέση συμπεριφορά.

Ενισχυτική Μάθηση: έννοιες, βασικές αρχές και ανάπτυξη προτύπου

Για την ανάπτυξη του Αυτογνωστικού Βοηθού Συστάσεων Οδήγησης (Self-Aware Driving Recommendation Assistant – SADRA) ακολουθήθηκε μία δομημένη διαδικασία. Αρχικά, η συνολική βάση δεδομένων χωρίστηκε στα δύο με βάση το μέσο προφίλ οδήγησης κάθε οδηγού. Συγκεκριμένα, η πρώτη βάση δεδομένων περιλαμβάνει τις διαδρομές όλων των οδηγών που ανήκουν στο «τυπικό/ασφαλές» προφίλ, ενώ η δεύτερη τις διαδρομές των οδηγών με μη ασφαλή μέση συμπεριφορά οδήγησης. Στο εξής, για συντομία, το μοντέλο Ενισχυτικής Μάθησης που αντιστοιχεί στους τυπικούς οδηγούς θα αναφέρεται ως SADRA – I, ενώ αυτό που αντιστοιχεί στους περισσότερο ριψοκίνδυνους οδηγούς ως SADRA – II.

Κάθε δομή Ενισχυτικής Μάθησης αποτελείται από τρία βασικά συστατικά: τις *καταστάσεις (s)* του συστήματος, τις πιθανές *δράσεις (a)* και τις *ανταμοιβές (r)*. Σε κάθε διακριτή χρονική στιγμή, ο κάθε πράκτορας παρατηρεί την τρέχουσα κατάσταση του περιβάλλοντός του και πραγματοποιεί την κατάλληλη ενέργεια από το σύνολο των πιθανών ενεργειών. Στη συνέχεια, ο πράκτορας λαμβάνει μια επιβράβευση, η οποία αντιστοιχεί στο κατά πόσο ήταν πετυχημένη ή αποτυχημένη η ενέργεια, σύμφωνα πάντα με τη συγκεκριμένη κατάσταση.

Στην παρούσα διατριβή, η κατάσταση του περιβάλλοντος ορίζεται μέσω ενός διανύσματος πέντε μεταβλητών που περιγράφουν τη συμπεριφορά οδήγησης του οδηγού κατά τη διάρκεια ενός ταξιδιού και περιλαμβάνουν τη μέση επιτάχυνση του ταξιδιού (a_{avg}), την επιτάχυνση που δεν ξεπέρασε ο οδηγός στο 90% του ταξιδιού (a_{90}), τη μέση επιβράδυνση (d_{avg}), την επιβράδυνση που δεν ξεπέρασε ο οδηγός στο 90% του ταξιδιού (d_{90}) και το ποσοστό του ταξιδιού που ο οδηγός οδηγούσε με ταχύτητα πάνω από το επιτρεπόμενο όριο ταχύτητας (*speeding*):

$$\mathbf{s} = \{a_{avg}, a_{90}, d_{avg}, d_{90}, speeding\}$$

Το παρόν σύστημα παροχής συστάσεων στοχεύει στο να βελτιώσει τη συμπεριφορά οδήγησης του κάθε χρήστη, ανεξάρτητα από τις εκάστοτε επικρατούσες κυκλοφοριακές συνθήκες. Γενικά, η επιλογή ταχύτητας κίνησης δεν είναι ανεξάρτητη από τη γεωμετρία της οδού που κινείται το όχημα και τις κυκλοφοριακές συνθήκες, όπως και η απόφαση για επιβράδυνση, η οποία εξαρτάται συνήθως από τη συμπεριφορά του προπορευόμενου οχήματος και τη σηματοδότηση. Συνεπώς, η παράμετρος που περιγράφει τη συμπεριφορά οδήγησης ενός οδηγού είναι η επιτάχυνση και εξαρτάται αποκλειστικά από την αντίληψή του και την προτίμησή του ανάμεσα σε ομαλή και απότομη επιτάχυνση. Πράγματι, και στην πρόσφατη βιβλιογραφία, η συμπεριφορά οδήγησης ενός οδηγού περιγράφεται συνήθως από το προφίλ επιτάχυνσής του.

Σε αυτή τη λογική, οι ενέργειες που προτείνονται από το σύστημα στον εκάστοτε οδηγό ανήκουν σε ένα συνεχές διάστημα τιμών και αφορούν στη μεταβολή (σε σχέση με την πιο πρόσφατη διαδρομή του) στη μέση επιτάχυνση και στην επιτάχυνση που δεν πρέπει ο οδηγός να ξεπερνά στο 90% των περιπτώσεων, π.χ. όταν πραγματοποιεί προσπέραση, εκτός και αν βρεθεί σε κατάσταση έκτακτης ανάγκης.

$$\boldsymbol{a} = \{da_{avg}, da_{90}\}$$

Στο εξής και για συντομία, η επιτάχυνση που δεν πρέπει ο οδηγός να ξεπεράσει στο 90% του ταξιδιού του θα αναφέρεται ως «μέγιστη επιτάχυνση».

Ένα πολύ σημαντικό κομμάτι του πλαισίου Ενισχυτικής Μάθησης είναι η συνάρτηση ανταμοιβής. Ο στόχος της είναι διττός: να αξιολογεί τόσο την παρούσα κατάσταση, όσο και τη μετάβαση μεταξύ διαδοχικών καταστάσεων. Στην προκειμένη περίπτωση, αξιολογεί τη συμπεριφορά οδήγησης σε κάθε διαδρομή, αλλά και τη μεταβολή της ανάμεσα σε διαδοχικές διαδρομές του ίδιου οδηγού. Στα πλαίσια της παρούσας διατριβής, αναπτύχθηκε μια ειδική συνάρτηση αξιολόγησης της συμπεριφοράς οδήγησης. Για κάθε διαδρομή υπολογίζεται ένα σκορ με βάση την απόκλισή της από το κέντρο της συστάδας του τυπικού προφίλ οδήγησης. Για τον υπολογισμό της απόκλισης χρησιμοποιήθηκε η απόσταση Mahalanobis. Η αξιολόγηση κάθε ταξιδιού δίνεται από την παρακάτω εξίσωση:

$$trip\ score_i = e^{-driving\ profile_i * \frac{\text{Ɱ}_{(i,moderate\ profile)}}{Q_{75}(\text{Ɱ})}}$$

όπου με i συμβολίζεται μια συγκεκριμένη διαδρομή, Ɱ είναι η απόσταση Mahalanobis και $Q_{75}(\text{Ɱ})$ είναι η τιμή που δεν ξεπερνά το 75% των αποστάσεων Mahalanobis.

Στη συνέχεια, υπολογίζεται η ανταμοιβή που αντιστοιχεί στη μετάβαση από τη μία διαδρομή στην άλλη, σύμφωνα με τον παρακάτω τύπο:

$$\boldsymbol{r} = trip\ score_{i+1}\left(1 + \frac{trip\ score_{i+1} - trip\ score_{i}}{100}\right)$$

Τελικά, αφού υπολογίστηκαν όλες οι απαραίτητες μεταβλητές για την ανάπτυξη του μοντέλου Ενισχυτικής Μάθησης (μοντέλο ή ελεγκτής ή πράκτορας RL), η βάση δεδομένων οργανώθηκε ως εξής:

(κατάσταση, δράση, ανταμοιβή, επόμενη κατάσταση)

Για κάθε μοναδικό οδηγό στο σύνολο δεδομένων, οι διαδρομές του ταξινομήθηκαν σε αύξουσα σειρά σύμφωνα με την ημερομηνία έναρξης κάθε διαδρομής. Τα δείγματα που χρησιμοποιήθηκαν για την εκπαίδευση του μοντέλου ήταν πλειάδες (tuples) διαδοχικών διαδρομών ενός συγκεκριμένου οδηγού μαζί με την αντίστοιχη ενέργεια και ανταμοιβή της μετάβασης από την πρώτη διαδρομή στην επόμενη. Πρέπει να σημειωθεί ότι για κάθε ξεχωριστό οδηγό στο σύνολο δεδομένων, η πρώτη διαδρομή του χρησιμοποιήθηκε μόνο ως "κατάσταση", ενώ η τελευταία διαδρομή του χρησιμοποιήθηκε μόνο ως "επόμενη κατάσταση". Μετά από αυτή τη διαδικασία προετοιμασίας δεδομένων, κατασκευάστηκαν 33.440 μοναδικά δείγματα δεδομένων για την εκπαίδευση του SADRA I και 119.817 μοναδικά δείγματα δεδομένων χρησιμοποιήθηκαν για τη διαδικασία εκπαίδευσης του SADRA II.

Οι ελεγκτές RL αναπτύχθηκαν με βάση τον αλγόριθμο Deep Deterministic Policy Gradient algorithm (DDPG), ο οποίος εφαρμόζει μια προσέγγιση "actor-critic" δηλαδή «ενέργειας και αξιολόγησης» για την εκμάθηση μιας πολιτικής και την παραγωγή των βέλτιστων ενεργειών. Έτσι, για κάθε ελεγκτή αναπτύσσονται δύο νευρωνικά δίκτυα που αντιπροσωπεύουν τη δράση (actor - μ) και την αξιολόγηση (critic - Q) αντίστοιχα. Τα νευρωνικά δίκτυα τόσο για το υποσύνολο των ασφαλών όσο και για το υποσύνολο των μη ασφαλών οδηγών εκπαιδεύτηκαν σύμφωνα με τη διαδικασία του παρακάτω αλγορίθμου.

DDPG Algorithm implementation

Initialize critic $Q(s,a|\theta^Q)$ and actor $\mu(s|\theta^\mu)$ networks using rewards as Q-values
Set the above as initial target networks (Q' and μ')
Split the sample into M minibatches
for minibatch=1, M **do**
  Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$
  Update critic by minimizing the loss: $L = \frac{1}{N}\sum_i \left(y_i - Q(s_i, a_i|\theta^Q)\right)^2$
  Update the actor policy using the sampled policy gradient:

$$\nabla_{\theta_\mu} J \approx \frac{1}{N}\sum_t [\nabla_a Q(s,a|\theta^Q)|_{s=s_t, a=\mu(s_t)}\ \nabla_{\theta_\mu}\mu(s|\theta^\mu)|_{s=s_t}]$$

  Update the target networks:

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$$
$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$$

end for

Η τελική δομή των δικτύων καθώς και οι τιμές των υπερπαραμέτρων προέκυψαν μετά από αξιολόγηση των πιθανών συνδυασμών. Συγκεκριμένα, εξετάστηκαν και συγκρίθηκαν όλοι οι πιθανοί συνδυασμοί δομών και παραμετροποίησης των δικτύων, σε ένα εύρος λογικών τιμών, προκειμένου να εντοπιστεί ο βέλτιστος. Οι παράμετροι που λήφθηκαν υπόψη είναι: αριθμός κρυφών στρωμάτων (*Number of hidden layers*), αριθμός νευρώνων (*Number of neurons per layer*) και ενεργοποίηση κάθε στρώματος (*Activation*), συνάρτηση βελτιστοποίησης (*Optimizer)* και ρυθμός μάθησης (*Learning rate*), μέγεθος δέσμης (*Batch size*) και αριθμός εποχών εκπαίδευσης (*Epochs*), όπως φαίνεται στον Πίνακα V.

Πίνακας V. Υπερπαράμετροι των δικτύων "critic" και "actor" για τα μοντέλα SADRA I και II

Hyperparameters	**Critic network**	**Actor network**
	SADRA I – ασφαλείς οδηγοί	
Number of hidden layers	6	3
Number of neurons per layer	(64,32,16,16,32,64,1)	(128,64,32,2)
Epochs	200(initial network:110)	200(initial network:110)
Batch size	150(initial network:150)	150(initial network:150)
Activation	ReLU	ReLU
Optimizer	Adam	Adam
Learning rate	0,001	0,0001
	SADRA II – μη-ασφαλείς οδηγοί	
Number of hidden layers	6	3
Number of neurons per layer	(32,16,8,8,16,32,1)	(128,64,32,2)
Epochs	170(initial network:170)	210(initial network:210)
Batch size	250(initial network:250)	250(initial network:100)
Activation	ReLU	ReLU
Optimizer	Adam	Adam
Learning rate	0,0001	0,0001

Σενάριο προσομοίωσης

Η ποσοτικοποίηση των επιπτώσεων της υιοθέτησης συστάσεων οδήγησης από όλους τους οδηγούς, στην κυκλοφορία, την οδική ασφάλεια και τις εκπομπές πραγματοποιήθηκε στο πλαίσιο ενός σεναρίου μικροσκοπικής προσομοίωσης (microscopic simulation) σε επίπεδο δικτύου. Για το σκοπό αυτό χρησιμοποιήθηκε το λογισμικό προσομοίωσης SUMO και το προεπιλεγμένο κυκλοφοριακό μοντέλο του Krauss, το οποίο είναι ένα μικροσκοπικό και συνεχές στο χώρο μοντέλο που βασίζεται στην ασφαλή ταχύτητα που σημαίνει ότι ο οδηγός του ακολουθούντος οχήματος υιοθετεί μια ασφαλή ταχύτητα που του επιτρέπει να προσαρμοστεί στην επιβράδυνση του προπορευόμενου οχήματος.

Η μελέτη περίπτωσης για τα πειράματα προσομοίωσης είναι το οδικό δίκτυο του δακτυλίου της Αθήνας, το οποίο αποτελείται από 1.293 κόμβους και 2.572 συνδέσμους. Το συνολικό μήκος του δικτύου είναι 348 χιλιόμετρα. Βάσει της βαθμονόμησης η οποία πραγματοποιήθηκε, στο δίκτυο κινούνται κατά την ώρα αιχμής 86.054 οχήματα. Από τους μετρητές προκύπτουν 1.393.634 μετρήσεις (97,47% των συνολικών μετρήσεων που εξήχθησαν από το πρόγραμμα προσομοίωσης Aimsun) και τιμή GEH κάτω από 5 (GEH < 5) για το 95,26%.

Δύο διαφορετικά σενάρια προσομοίωσης σχεδιάστηκαν που αντιστοιχούν στη ζήτηση του οδικού δικτύου της Αθήνας κατά την πρωινή ώρα αιχμής (8:00 - 9:00 π.μ.). Αρχικά, προσομοιώνονται οι αρχικές συνθήκες του δικτύου προκειμένου να εκτιμηθεί η απόδοση της κυκλοφορίας όταν τα οχήματα κινούνται σύμφωνα με τα χαρακτηριστικά που διέπουν τα έξι προφίλ οδήγησης που προσδιορίστηκαν. Προκειμένου να διασφαλιστεί η αξιοπιστία των αποτελεσμάτων, η προσομοίωση πραγματοποιήθηκε σε 10 επαναλήψεις με δέκα διαφορετικούς αριθμούς εκκίνησης. Η στοχαστικότητα αποτελεί σημαντική πτυχή της αναπαραγωγής της πραγματικότητας σε ένα σενάριο προσομοίωσης, καθώς προσθέτει τυχαιότητα στις κατανομές των διαφορετικών πτυχών της προσομοίωσης (π.χ. κατανομές διαδρομών, κατανομές τύπων οχημάτων).

Στη συνέχεια, δημιουργήθηκαν συστάσεις οδήγησης για κάθε εξυπηρετούμενο όχημα με βάση τον τρόπο με τον οποίο κάθε όχημα πραγματοποίησε τη διαδρομή του στο πρώτο σενάριο προσομοίωσης. Οι συστάσεις παρήχθησαν από τους αντίστοιχους ελεγκτές RL χρησιμοποιώντας ως είσοδο την κατάσταση της διαδρομής (μέση επιτάχυνση, 90% εκατοστημόριο επιτάχυνσης, μέση επιβράδυνση, 90% εκατοστημόριο επιβράδυνσης, ποσοστό επιτάχυνσης) και ως έξοδο τη βέλτιστη μεταβολή της μέγιστης επιτάχυνσης. Θα πρέπει να τονιστεί εδώ ότι παρόλο που οι ελεγκτές RL που αναπτύχθηκαν παράγουν ένα δισδιάστατο διάνυσμα που περιλαμβάνει μεταβολές τόσο στη μέση όσο και στη μέγιστη επιτάχυνση, στο πλαίσιο της προσομοίωσης αξιοποιήθηκε μόνο η μέγιστη επιτάχυνση, καθώς το μοντέλο Krauss λαμβάνει υπόψη μόνο τις μέγιστες τιμές της επιτάχυνσης και της επιβράδυνσης.

Τέλος, ένα δεύτερο σενάριο προσομοίωσης εκτελέστηκε, όπου τα οχήματα που είχαν εξυπηρετηθεί προηγουμένως ακολουθούν τις προτεινόμενες συστάσεις, δηλαδή μια εναλλαγή της μέγιστης επιτάχυνσής τους, ενώ η υπόλοιπη κυκλοφορία ακολουθεί την κατανομή μεταξύ των έξι προφίλ οδήγησης.

Η συμπεριφορά που συνεπάγεται κάθε προφίλ οδήγησης προσομοιώθηκε μέσω της προσαρμογής του αντίστοιχου κυκλοφοριακού μοντέλου. Σε αυτή την περίπτωση χρησιμοποιήθηκε το μοντέλο ακολουθούντος οχήματος Krauss το οποίο μπορεί να παραμετροποιηθεί από έναν αριθμό παραμέτρων: τη μέγιστη επιτάχυνση του οχήματος (accel), τη μέγιστη επιβράδυνση του οχήματος (decel), τη μέγιστη ταχύτητα του οχήματος (maxSpeed), τη μέγιστη φυσικά δυνατή επιβράδυνση του οχήματος (emergencyDecel) και τον αναμενόμενο πολλαπλασιαστή για τα όρια ταχύτητας της λωρίδας (speedFactor). Αρχικά, η τρέχουσα (αρχική) κατάσταση της οδικής κυκλοφορίας προσομοιώνεται στο SUMO χρησιμοποιώντας τα έξι καθορισμένα προφίλ οδήγησης, οι παράμετροι των οποίων εισήχθησαν στο μοντέλο Krauss για διαφορετικούς τύπους οχημάτων, όπως φαίνεται στον Πίνακα VI.

Πίνακας VI. Παραμετροποίηση του μοντέλου ακολουθούντος οχήματος για κάθε τύπο οχήματος

Τύποι οχήματος (προφίλ οδήγησης)	**Παράμετροι μοντέλου ακολουθούντος οχήματος**				
	accel (m/s^2)	**decel** (m/s^2)	**emergencyDecel** (m/s^2)	**maxSpeed** (km/h)	**speedFactor** (mean, min, max)
Τυπική οδήγηση	2,519	-2,942	-5,909	64,51	(0,029, 0, 0,168)
Επιθετική	3,817	-4,483	-18,083	66,93	(0,033, 0, 0,151)
Ανάληψη ρίσκου	2,392	-2,824	-5,328	100,28	(0,306, 0,1627, 0,96)
Απόσπαση προσοχής	2,601	-2,990	-5,112	67,38	(0,057, 0, 0,631)
Επιθετική με ανάληψη ρίσκου	3,944	-4,825	-25,884	100,8	(0,269, 0,147, 0,907)
Επιθετική με απόσπαση προσοχής	3,939	-4,553	-10,845	71,99	(0,062, 0, 0,744)

Για την αρχική κατάσταση του δικτύου, οι έξι διαφορετικοί τύποι οχημάτων δημιουργήθηκαν σε ένα αρχείο διαδρομών (route file), με την αντίστοιχη παραμετροποίηση του μοντέλου ακολουθούντος οχήματος. Σε μία ώρα προσομοίωσης για την πρωινή αιχμή, περίπου το 58% της συνολικής ζήτησης εισήχθη στο δίκτυο και το 28% των οχημάτων ολοκλήρωσε τη διαδρομή του μέσα σε αυτό το χρονικό διάστημα.

Στη συνέχεια, για κάθε όχημα που έφτασε στον προορισμό του εκτιμήθηκαν οι ακόλουθες παράμετροι για κάθε διαδρομή:

- μέση επιτάχυνση
- 90% εκατοστημόριο επιτάχυνσης
- μέση επιβράδυνση
- 90% εκατοστημόριο επιβράδυνσης
- ποσοστό επιτάχυνσης

Αυτά τα χαρακτηριστικά οδήγησης χρησιμοποιήθηκαν ως δεδομένα εισόδου στους ελεγκτές RL, οι οποίοι προτείνουν τη βέλτιστη ενέργεια για κάθε διαδρομή. Για την εκτέλεση του δεύτερου σεναρίου της προσομοίωσης χρησιμοποιήθηκαν τα ίδια ακριβώς οχήματα, τα οποία ακολουθούν τις ίδιες ακριβώς διαδρομές στο ίδιο οδικό δίκτυο, προκειμένου να εκτιμηθούν οι επιπτώσεις της παροχής προσωποποιημένων συστάσεων. Οι συστάσεις για αλλαγή της μέγιστης επιτάχυνσης εισήχθησαν ως τροποποίηση της αντίστοιχης παραμέτρου του μοντέλου ακολουθούντος οχήματος. Η υιοθέτηση αυτής της προσέγγισης επέτρεψε την πρακτική εφαρμογή της διαδικασίας εφαρμογής των συστάσεων με άμεσο έλεγχο των αποτελεσμάτων.

Και σε αυτή την περίπτωση, του δεύτερου σεναρίου της προσομοίωσης, πραγματοποιήθηκαν 10 επαναλήψεις με τις ίδιες τιμές εκκίνησης όπως και προηγουμένως, ώστε να διασφαλιστεί η αξιοπιστία των αποτελεσμάτων. Τα ευρήματα έδειξαν ότι σε μία ώρα προσομοίωσης εξυπηρετήθηκε κατά μέσο όρο το 57% της ζήτησης, ενώ το αντίστοιχο ποσοστό των οχημάτων που ολοκλήρωσαν τη διαδρομή τους μειώθηκε κατά 1% σε σχέση με τις αρχικές συνθήκες.

Αξιολόγηση επιπτώσεων

Η αξιολόγηση των επιπτώσεων του προτεινόμενου συστήματος πραγματοποιείται μέσω μιας προσέγγισης "πριν και μετά" όπως περιεγράφηκε παραπάνω. Συγκεκριμένα, και για τα δύο σενάρια προσομοίωσης εκτιμήθηκαν οι βασικοί δείκτες απόδοσης της κυκλοφορίας, της ασφάλειας και των περιβαλλοντικών συνθηκών και αξιολογήθηκαν συγκριτικά, ώστε να ποσοτικοποιηθούν οι συνολικές επιπτώσεις της υιοθέτησης εξατομικευμένων συστάσεων που βελτιώνουν τη συμπεριφορά οδήγησης κάθε ατόμου. Οι βασικοί δείκτες απόδοσης που χρησιμοποιήθηκαν στην ανάλυση για κάθε πτυχή του δικτύου παρουσιάζονται στον Πίνακα VII.

Πίνακας VII. Βασικοί δείκτες απόδοσης για τις διάφορες πτυχές του δικτύου

Κυκλοφορία	Ασφάλεια	Περιβάλλον
Εξυπηρετούμενη ζήτηση	Συνολικές πιθανές εμπλοκές	Συνολικές εκπομπές ανά τύπο ρύπου (CO_2, CO, PMx, NOx)
Μακροσκοπικό θεμελιώδες διάγραμμα κυκλοφορίας	Συνολικές μετόπισθεν εμπλοκές	Εκπομπές ανά όχημα
	Πιθανές εμπλοκές ανά όχημα	

Η εκτίμηση των βασικών δεικτών απόδοσης που αφορούν στην κυκλοφορία έγινε με βάση τα αποτελέσματα της προσομοίωσης, τα οποία περιλάμβαναν τον αριθμό των εισαγόμενων και εξυπηρετούμενων οχημάτων, καθώς και πληροφορίες για τα τρία θεμελιώδη στοιχεία της θεωρίας της κυκλοφοριακής ροής (ροή, ταχύτητα και πυκνότητα). Αντί να χρησιμοποιηθούν συγκεντρωτικές μετρήσεις των θεμελιωδών μεταβλητών, κατασκευάστηκαν τα μακροσκοπικά θεμελιώδη διαγράμματα (Macroscopic Fundamental Diagrams – MFDs) και εξήχθησαν σημαντικά αποτελέσματα σχετικά με τις διαφορές στις επιδόσεις του δικτύου πριν και μετά την εφαρμογή του συστήματος συστάσεων. Η εκτίμηση των επιβλαβών ατμοσφαιρικών ρύπων βασίζεται στο μοντέλο εκπομπών που είναι ήδη ενσωματωμένο στο SUMO, το μοντέλο PHEMlight. Το PHEMlight είναι μια απλουστευμένη έκδοση του PHEM (Passenger car and Heavy-duty Emission Model), ενός πλήρους μοντέλου εκπομπών οχημάτων που έχει αναπτυχθεί στην Ευρώπη από το 1999 και βασίζεται σε εκτεταμένες μετρήσεις εκπομπών σε οχήματα όπως επιβατικά αυτοκίνητα, ελαφρά οχήματα και αστικά λεωφορεία. Η εκτίμηση των εμπλοκών που αποτελούν δείκτη οδικής ασφάλειας βασίζεται στο εργαλείο SSAM, το οποίο υπολογίζει υποκατάστατα μέτρα ασφάλειας για κάθε εμπλοκή που εντοπίζεται στα δεδομένα τροχιάς, και στη συνέχεια υπολογίζει τα στατιστικά χαρακτηριστικά (μέση τιμή, μέγιστη τιμή κ.λπ.) κάθε υποκατάστατου μέτρου (surrogate measure).

Αποτελέσματα: Συστάσεις οδήγησης

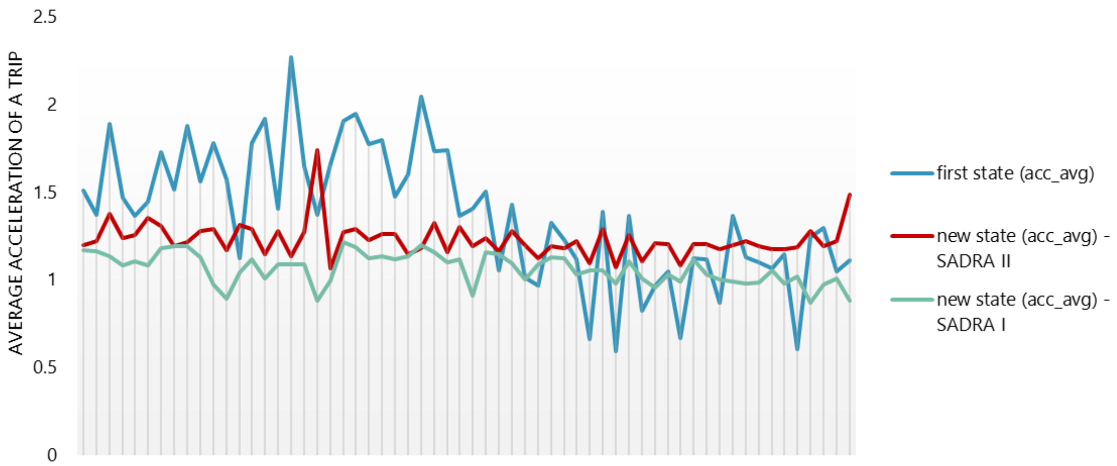
Οι δύο εκδόσεις του εκπαιδευμένου αλγορίθμου DDPG χρησιμοποιήθηκαν για την παραγωγή συστάσεων οδήγησης για δύο κατηγορίες οδηγών: τυπικοί οδηγοί που παρουσιάζουν μέτρια-ασφαλή συμπεριφορά (SADRA I) και μη-ασφαλείς οδηγοί που εναλλάσσουν τη συμπεριφορά τους μεταξύ διαφόρων ανασφαλών συνηθειών οδήγησης (SADRA II). Οι συστάσεις έχουν τη μορφή αλλαγών στον τρόπο οδήγησης που αναφέρονται στις βέλτιστες ενέργειες οδήγησης που μπορεί να υιοθετήσει ο συγκεκριμένος οδηγός προκειμένου να βελτιώσει την οδήγησή του με βάση την τρέχουσα συμπεριφορά του.

Η σύγκριση των αποτελεσμάτων των δύο ελεγκτών αποκάλυψε ότι και οι δύο έχουν εκπαιδευτεί να παράγουν συστάσεις που φέρνουν τους οδηγούς πιο κοντά στη μέση ασφαλή συμπεριφορά ενός τυπικού οδηγού, ο οποίος έχει μέση επιτάχυνση ίση με 1,137 m/s^2 και μέγιστη επιτάχυνση ίση με 2,503 m/s^2. Με βάση τα ενδεικτικά παραδείγματα του παρακάτω πίνακα (Πίνακας VIII), η μέση συνιστώμενη μέση επιτάχυνση εκτιμήθηκε σε 1,145 m/s^2, ενώ η μέση τιμή των προτεινόμενων μέγιστων επιταχύνσεων ήταν 2,507 m/s^2 αντίστοιχα. Επομένως, μπορεί κανείς να συμπεράνει ότι η καθολική εφαρμογή του προτεινόμενου συστήματος συστάσεων θα οδηγούσε στην εναρμόνιση των προφίλ επιτάχυνσης για ολόκληρο το στόλο οχημάτων.

Πίνακας VIII. Παραδείγματα εισόδου και εξόδου RL και των παραγόμενων συστάσεων

Προφίλ και βαθμολόγηση		RL είσοδος					RL έξοδος		Σύσταση οδήγησης	
Προφίλ οδήγησης	**Σκορ**	**Μέση επιτάχυνση**	**Μέση επιβράδυνση**	**Επιτάχυνση Q90**	**Επιβράδυνση Q90**	**Ποσοστό οδήγησης πάνω από το όριο**	**ενέργεια_1 (Q90)**	**ενέργεια_2 (μέση)**	**νέα επιτάχυνση Q90**	**νέα μέση επιτάχυνση**
Επιθετικότητα με απόσπαση προσοχής	9	1.512	-1.58	3.600	-3.72	0.0090	-0.99	-0.31	2.610	1.197
Επιθετικότητα με απόσπαση προσοχής	2	1.888	-1.86	3.336	-3.95	0.0660	-0.29	-0.51	3.046	1.377
Επιθετικότητα με ανάληψη ρίσκου	6	1.779	-1.63	3.432	-3.82	0.3235	-0.65	-0.49	2.778	1.288
Επιθετικότητα με ανάληψη ρίσκου	2	1.574	-1.73	3.480	-4.20	0.5693	-0.96	-0.40	2.519	1.171
Επιθετική οδήγηση	2	1.905	-1.68	3.600	-3.18	0.0000	-0.77	-0.63	2.834	1.275
Επιθετική οδήγηση	17	1.948	-2.57	3.768	-5.69	0.0822	-1.10	-0.66	2.669	1.289
Επιθετική οδήγηση	10	1.774	-2.17	4.368	-4.08	0.0000	-1.69	-0.54	2.676	1.230
Απόσπαση προσοχής	2	1.405	-1.33	4.104	-2.88	0.0133	-1.38	-0.21	2.727	1.193
Απόσπαση προσοχής	24	1.504	-1.52	3.324	-3.36	0.0000	-0.64	-0.27	2.679	1.236
Απόσπαση προσοχής	35	1.054	-1.08	2.489	-2.67	0.0000	0.06	0.11	2.548	1.166
Τυπική οδήγηση	73	1.386	-1.59	2.760	-3.28	0.0226	-0.46	-0.33	2.299	1.055
Τυπική οδήγηση	58	0.592	-0.92	1.560	-2.81	0.0294	0.56	0.39	2.123	0.980
Τυπική οδήγηση	86	1.366	-1.43	2.880	-3.36	0.0000	-0.45	-0.26	2.426	1.106
Τυπική οδήγηση	90	0.826	-1.03	1.984	-2.43	0.0000	0.26	0.18	2.240	1.006
Τυπική οδήγηση	67	0.967	-1.43	2.040	-3.67	0.0674	0.05	-0.01	2.091	0.954
Τυπική οδήγηση	78	1.047	-1.23	2.070	-2.33	0.0000	0.21	-0.02	2.284	1.030
Τυπική οδήγηση	77	0.666	-0.69	1.440	-1.68	0.0367	0.76	0.32	2.195	0.989
Τυπική οδήγηση	94	1.124	-1.24	2.568	-2.64	0.0000	-0.06	-0.01	2.511	1.118
Τυπική οδήγηση	58	0.871	-1.24	1.920	-3.56	0.1667	0.27	0.13	2.187	1.003
Ανάληψη ρίσκου	52	1.367	-1.36	3.093	-3.06	0.2392	-0.49	-0.17	2.602	1.196
Ανάληψη ρίσκου	33	1.098	-1.43	2.280	-3.24	0.2857	0.32	0.09	2.600	1.192

Το Σχήμα III παρουσιάζει ενδεικτικά παραδείγματα των συστάσεων που παράγονται από τους δύο ελεγκτές με δεδομένη την ίδια είσοδο (πρώτη κατάσταση). Τα ευρήματα έδειξαν ότι, παρόλο που οι συστάσεις του ελεγκτή που αφορά στους μη ασφαλείς οδηγούς (SADRA II) οδηγούν σε σημαντικά χαμηλότερες μέσες επιταχύνσεις για την επόμενη διαδρομή (επόμενη κατάσταση) σε σύγκριση με την προηγούμενη διαδρομή (αρχική κατάσταση), διατηρούν σημαντική απόσταση προς τα πάνω σε σχέση με τις αντίστοιχες συστάσεις που παράγονται από τον ελεγκτή των τυπικών οδηγών (SADRA I). Παρ' όλα αυτά, πρέπει να σημειωθεί ότι και οι δύο ελεγκτές οδηγούν σε ομαλότερο προφίλ επιτάχυνσης για το σύνολο των οχημάτων που απαρτίζουν την κυκλοφορία.

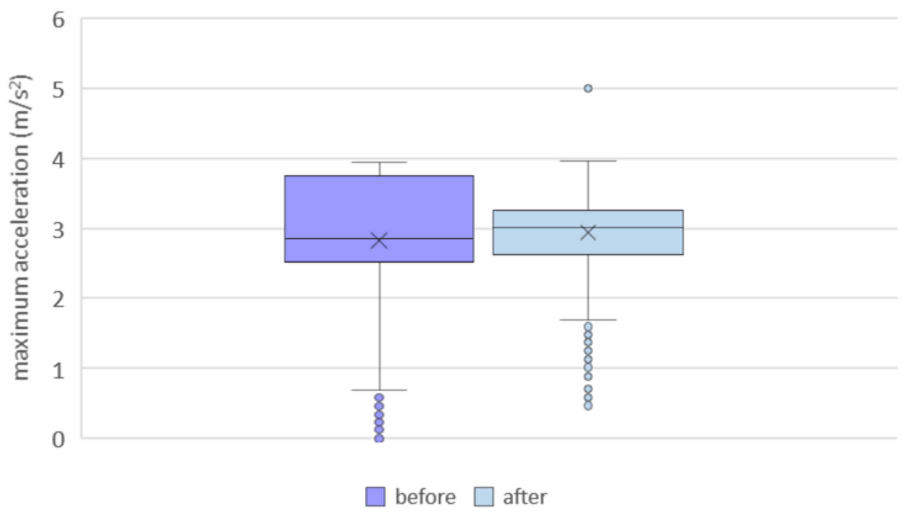


Σχήμα III. Σύγκριση της μέσης επιτάχυνσης για τη νέα κατάσταση όπως προέκυψε από το SADRA I (κόκκινο) και SADRA II (πράσινο) με βάση την αρχική κατάσταση (μπλε).

Αποτελέσματα: Προσομοίωση και αξιολόγηση επιπτώσεων

Η ποσοτικοποίηση των επιπτώσεων της εφαρμογής του προτεινόμενου συστήματος συστάσεων και, κατά συνέπεια, της υιοθέτησης μιας βελτιωμένης συμπεριφοράς οδήγησης από όλους τους οδηγούς έχει μεγάλη σημασία, τόσο για τους ερευνητές, όσο και για τους επαγγελματίες και μπορεί να οδηγήσει σε σημαντικά συμπεράσματα σχετικά με τη χρησιμότητα της βελτίωσης της ατομικής συμπεριφοράς οδήγησης. Η αξιολόγηση του συστήματος συστάσεων πραγματοποιείται με τη χρήση συγκεκριμένων βασικών δεικτών επιδόσεων που αντιστοιχούν σε τρεις τομείς ενδιαφέροντος: κυκλοφορία, ασφάλεια και εκπομπές. Κάθε ένα από τα σενάρια προσομοίωσης έγινε σε 10 επαναλήψεις για να εξασφαλιστεί η εγκυρότητα και η αξιοπιστία των αποτελεσμάτων. Συνολικά, ο ελεγκτής SADRA I χρησιμοποιήθηκε για την παραγωγή συστάσεων για το 43% των οχημάτων, ενώ τα υπόλοιπα οχήματα ακολούθησαν τις συστάσεις που παρήγαγε το SADRA II.

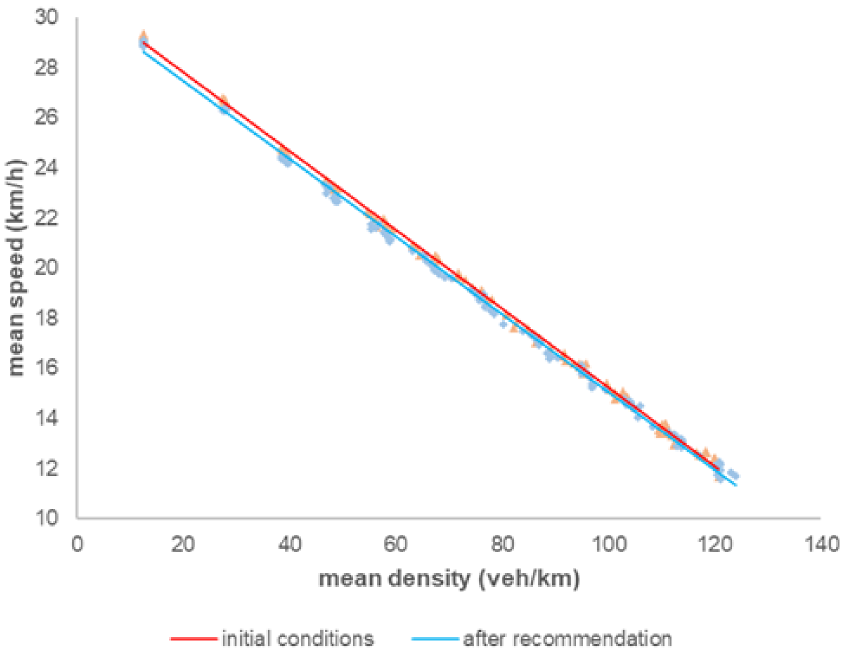
Όλες οι επαναλήψεις του ίδιου σεναρίου προσομοίωσης παρουσιάζουν αντίστοιχα αποτελέσματα σχετικά με τα εξυπηρετούμενα οχήματα, τα οποία μειώνονται ελαφρώς μετά την εφαρμογή του συστήματος συστάσεων. Κατά μέσο όρο, εξυπηρετήθηκαν 2,9% λιγότερα οχήματα με βάση τα αποτελέσματα του δεύτερου σεναρίου της προσομοίωσης. Ωστόσο, τα αποτελέσματα του στατιστικού ελέγχου υποθέσεων t-test έδειξαν ότι δεν υπάρχουν σημαντικές διαφορές μεταξύ των μέσων όρων των εξυπηρετούμενων οχημάτων πριν και μετά τις συστάσεις σε διάστημα εμπιστοσύνης 95%. Η εφαρμογή του συστήματος εξατομικευμένων συστάσεων είχε σημαντική επιρροή στη μέγιστη επιτάχυνση των οχημάτων, όπως φαίνεται στο Σχήμα IV. Όταν όλα τα οχήματα ακολούθησαν τις προτάσεις που παρήγαγαν οι δύο ελεγκτές RL, η μέση τιμή της μέγιστης επιτάχυνσης αυξήθηκε ελάχιστα από 2,83 m/s^2 σε 2,96 m/s^2, κυρίως επειδή η πλειονότητα των οχημάτων που αρχικώς είχαν μια πολύ μικρή μέγιστη επιτάχυνση, η οποία ήταν πολύ χαμηλότερη από την αντίστοιχη επιτάχυνση της "μέτριας/τυπικής" συμπεριφοράς, τους προτάθηκε να αυξήσουν ελαφρώς την επιτάχυνσή τους. Ωστόσο, η μείωση του εύρους των τιμών της επιτάχυνσης είναι εμφανής μετά τις συστάσεις, γεγονός που υποδηλώνει την εναρμόνιση των προφίλ επιτάχυνσης όλων των οχημάτων στην προσομοίωση. Τέλος, η μέγιστη τιμή των παρατηρούμενων μέγιστων επιταχύνσεων παρέμεινε στο ίδιο επίπεδο των 3,94 m/s^2 μετά την εφαρμογή του προτεινόμενου συστήματος.



Σχήμα IV. Boxplot των τιμών της μέγιστης επιτάχυνσης πριν και μετά τις συστάσεις

Οι διαφορές που παρατηρούνται στο μέγεθος της μέσης ταχύτητας είναι ελάχιστες, καθώς και στις δύο περιπτώσεις τα οχήματα υιοθετούν μέση ταχύτητα περίπου 25 km/h, ενώ η μέγιστη μέση ταχύτητα που παρατηρείται είναι περίπου 55 km/h.

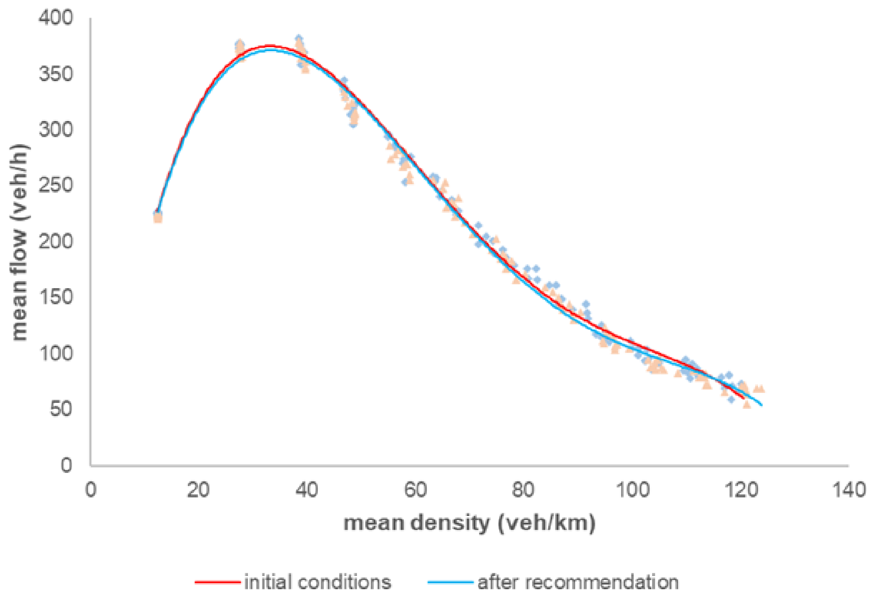
Οι μεταβολές της ταχύτητας των οχημάτων είχαν ως αποτέλεσμα μεταβολές των ιδιοτήτων των υπόλοιπων κυκλοφοριακών μεγεθών, δηλαδή της ροής και της πυκνότητας. Προκειμένου να αποδοθεί μια λεπτομερής γραφική απεικόνιση των σχέσεων αυτών των μεγεθών για τις αρχικές συνθήκες καθώς και για τις συνθήκες που προέκυψαν μετά τις συστάσεις, υπολογίστηκαν τα μακροσκοπικά θεμελιώδη διαγράμματα. Και τα τρία θεμελιώδη διαγράμματα (Σχήματα V - VII) δείχνουν τις σχέσεις μεταξύ των της ροής κυκλοφορίας, δηλαδή της μέσης ροής οχημάτων, της μέσης πυκνότητας και της μέσης ταχύτητας, όπως προέκυψαν από την προσομοίωση με βάση τις συγκεντρωτικές μετρήσεις όλων των συνδέσμων και για τις 10 επαναλήψεις. Τα αποτελέσματα δείχνουν ότι η υιοθέτηση των προτάσεων οδήγησης, αν και οδηγεί σε ασφαλέστερη και λιγότερο επιθετική συμπεριφορά οδήγησης για κάθε άτομο, δε βελτιώνει την απόδοση του οδικού δικτύου. Πιο συγκεκριμένα, η αυτό-βελτίωση είναι εμφανής από τις χαμηλότερες μέσες τιμές πυκνότητας, οι οποίες υποδηλώνουν ότι τα οχήματα διατηρούν μεγαλύτερες αποστάσεις από τα προπορευόμενα οχήματα. Επιπλέον, παρατηρούνται χαμηλότερες ταχύτητες μετά την προσαρμογή των επιταχύνσεων, με τη διαφορά από τις αρχικές συνθήκες να είναι πιο σημαντική στην περίπτωση κορεσμένης ροής του δικτύου (Σχήμα V).



Σχήμα V. Θεμελιώδες διάγραμμα ταχύτητας-πυκνότητας πριν (κόκκινο) και μετά (μπλε) τις συστάσεις

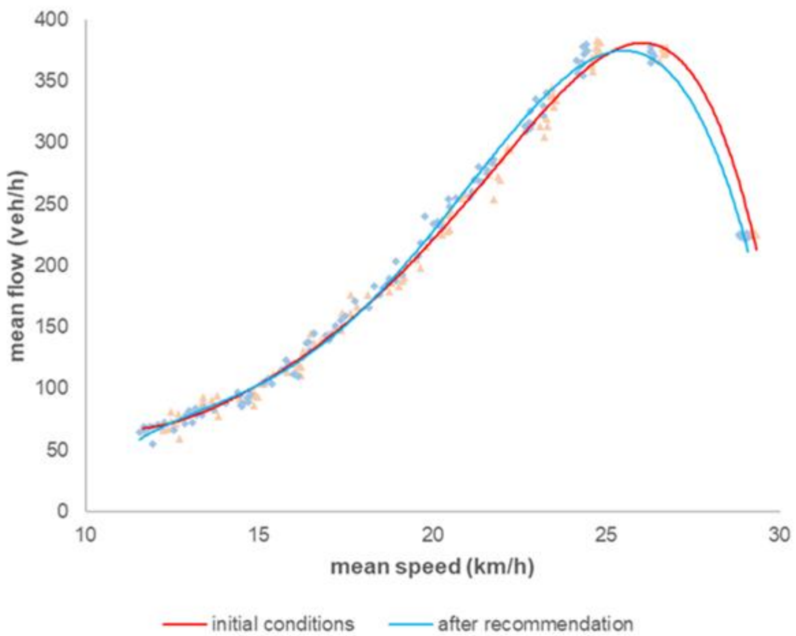
Η ασφάλεια οδήγησης σε ατομικό επίπεδο αυξάνεται, αλλά οι επιπτώσεις στις συνθήκες κυκλοφορίας δεν είναι εξίσου θετικές. Τα οχήματα που κινούνται με χαμηλότερες ταχύτητες και με μικρότερη πυκνότητα επιδεινώνουν τις συνθήκες κυκλοφοριακής ροής, καθώς εξυπηρετούνται λιγότερα οχήματα ανά μονάδα χρόνου σε σύγκριση με τις αρχικές συνθήκες. Ωστόσο, αυτή η μείωση της μέσης ροής μπορεί να θεωρηθεί αποδεκτή εάν αξιολογηθεί σε συνδυασμό με τις θετικές επιπτώσεις στην οδική ασφάλεια. Ωστόσο, με βάση τα ευρήματα της παρούσας έρευνας, δεν μπορεί σε καμία περίπτωση να εξαχθεί το συμπέρασμα ότι η βελτίωση της προσωπικής συμπεριφοράς οδήγησης συνδέεται με σημαντική βελτίωση των συνθηκών κυκλοφορίας και, ως εκ τούτου, η επιβολή μέτρων που βελτιώνουν τον τρόπο οδήγησης σε προσωπικό επίπεδο, όπως η αύξηση της αυτογνωσίας σε σχέση με την ατομική οδηγική ασφάλεια και τις συνέπειές της, δεν μπορεί να θεωρηθεί ως βασικό μέτρο για τη διαχείριση της κυκλοφορίας.

Το θεμελιώδες διάγραμμα ροής-πυκνότητας απεικονίζει μια ομοιομορφία μεταξύ των αρχικών και των τελικών συνθηκών, αν και παρατηρούνται ορισμένες μικρές διαφορές όσον αφορά την απόλυτη τιμή της ροής κορεσμού (Σχήμα VI). Συγκεκριμένα, για την τιμή της κρίσιμης πυκνότητας, η οποία εκτιμήθηκε 33,1 veh/km, οι τιμές της κυκλοφοριακής ροής είναι 360 veh/h και 358 veh/h για τις αρχικές συνθήκες και μετά τις συστάσεις αντίστοιχα.



Σχήμα VI. Θεμελιώδες διάγραμμα ροής-πυκνότητας πριν (κόκκινο) και μετά (μπλε) τις συστάσεις

Το θεμελιώδες διάγραμμα ταχύτητας-ροής χρησιμοποιείται για τον προσδιορισμό της ταχύτητας στην οποία εμφανίζεται η βέλτιστη ροή. Για τις αρχικές συνθήκες του οδικού δικτύου, η βέλτιστη ροή εμφανίζεται όταν τα οχήματα κινούνται με 26,1 km/h, ενώ η αντίστοιχη ταχύτητα μετά την εφαρμογή των συστάσεων μειώνεται κατά 3,4% με την απόλυτη τιμή της να εκτιμάται στα 25,2 km/h (Σχήμα VII).



Σχήμα VII. Θεμελιώδες διάγραμμα ροής-ταχύτητας πριν (κόκκινο) και μετά (μπλε) τις συστάσεις

Εκτός από την απόδοση του δικτύου, ένας άλλος βασικός δείκτης απόδοσης είναι η οδική ασφάλεια. Η αξιολόγηση των παραγόμενων συστάσεων όσον αφορά την ασφάλεια πραγματοποιήθηκε με τον υπολογισμό του αριθμού των εμπλοκών που σημειώθηκαν μεταξύ των οχημάτων κατά τη διάρκεια της προσομοίωσης. Στον Πίνακα IX παρουσιάζεται ο αριθμός των εμπλοκών που παρατηρήθηκαν για το σύνολο της κυκλοφορίας πριν και μετά τις συστάσεις. Υπάρχουν τρεις τύποι εμπλοκών που μπορούν να εντοπιστούν από τις τροχιές των οχημάτων, οι οποίοι είναι οι διασταυρώσεις οχημάτων (crossings), οι μετόπισθεν εμπλοκές (rear-ends) και οι αλλαγές λωρίδας (lane changes). Εδώ, δίνεται ιδιαίτερη έμφαση στις μετόπισθεν εμπλοκές, δεδομένου ότι οι προτεινόμενες συστάσεις επηρεάζουν μόνο τη συμπεριφορά του κάθε οδηγού σε σχέση με τον τρόπου που προσαρμόζει την οδήγησή του με βάση το προπορευόμενο όχημα (car-following behavior).

Πίνακας IX. Δείκτες απόδοσης ασφάλειας για το δίκτυο της Αθήνας πριν και μετά τις συστάσεις

	Αρχικές συνθήκες	Μετά τις συστάσεις [% μεταβολή]
Οχήματα που εξυπηρετήθηκαν	23.990	23.302
(σε μία ώρα προσομοίωσης)	*(27,88% της ζήτησης)*	*(27,08% της ζήτησης)*
Συνολικός αριθμός εμπλοκών	2,86 εμπλοκές/όχημα	2,75 εμπλοκές/όχημα [-4,2%]
Μετόπισθεν εμπλοκές	2,01 μετόπισθεν εμπλοκές/όχημα	1,90 μετόπισθεν εμπλοκές/όχημα [-5,5%]

Παρατηρήθηκε μείωση κατά 4,2% του συνολικού αριθμού των εμπλοκών όταν τα οχήματα ακολουθούσαν τις αντίστοιχες συστάσεις οδήγησης, ενώ το αντίστοιχο ποσοστό μείωσης των μετόπισθεν εμπλοκών είναι 5,5%. Αν και τα ποσοστά αυτά μπορεί να μην φαίνονται πολύ υψηλά, ο απόλυτος αριθμός συγκρούσεων που υπολογίστηκε μετά την εφαρμογή των συστάσεων μειώνεται σημαντικά κατά περίπου 6.000 συγκρούσεις για τη μία ώρα προσομοίωσης. Οι μετόπισθεν εμπλοκές αποτελούν περίπου το 33% του συνολικού αριθμού εμπλοκών, γεγονός που υποδηλώνει ότι κάθε οδηγός εμπλέκεται σε όλα τα διαφορετικά είδη εμπλοκών κατά τη διάρκεια της οδήγησης.

Τέλος, παρέχονται ορισμένα ενδεικτικά αποτελέσματα σχετικά με τις επιδράσεις του προτεινόμενου συστήματος συστάσεων στις εκπομπές ρύπων. Ο σχετικός δείκτης απόδοσης είναι το επίπεδο εκπομπών για όλα τα διαφορετικά είδη ατμοσφαιρικών ρύπων, δηλαδή διοξείδιο του άνθρακα (CO_2), μονοξείδιο του άνθρακα (CO), σωματιδιακή ύλη (PM_x) και οξείδια του αζώτου (NO_x). Παρατηρείται σημαντική μείωση σε όλες τις κατηγορίες εκπομπών σε σύγκριση με τις αρχικές συνθήκες του δικτύου, όπως φαίνεται στον Πίνακα X. Τα αποτελέσματα έδειξαν ότι η εξομάλυνση του προφίλ επιτάχυνσης για το σύνολο της κυκλοφορίας οδήγησε σε ελαφρώς μειωμένες εκπομπές ανά όχημα. Συγκεκριμένα, η μείωση σε όλες τις κατηγορίες εκπομπών εκτιμάται ως εξής: 2,5% για το CO_2, 0,3% για το CO, 1,3% για τα PM_x και 3,3% για τα NO_x. Πρέπει να σημειωθεί ότι αυτή η βελτίωση των περιβαλλοντικών συνθηκών είναι πολύ σημαντική, δεδομένου ότι το προτεινόμενο σύστημα συστάσεων είχε θετική επίδραση στις εκπομπές παρά το γεγονός ότι ο αλγόριθμος παραγωγής των συστάσεων δεν είχε εκπαιδευτεί προς αυτή την κατεύθυνση.

Πίνακας X. Μεταβολές στις εκπομπές αερίων πριν και μετά τις συστάσεις

Emissions	Αρχικές συνθήκες	Μετά τις συστάσεις [% μεταβολή]
CO_2	0,704 kg/όχημα	0,686 kg/ όχημα [-2.5%]
CO	0,027 kg/ όχημα	0,026 kg/ όχημα [-0.3%]
PM_x	0,0133 g/ όχημα	0,0131 g/ όχημα [-1.3%]
NO_x	0,296 g/ όχημα	0,287 g/ όχημα [-3.3%]

Συμπεράσματα και συνεισφορά διατριβής

Τα κύρια συμπεράσματα της διατριβής μπορούν να συνοψιστούν στα ακόλουθα σημεία:

- Μια προσέγγιση ομαδοποίησης δύο επιπέδων μπορεί να προσφέρει σημαντική γνώση σχετικά με τα χαρακτηριστικά που διέπουν την επιθετικότητα κατά τη διάρκεια της οδήγησης και μπορεί να αξιοποιηθεί περαιτέρω για τη διάκριση ασφαλών και μη ασφαλών προτύπων οδήγησης.
- Έξι διακριτά προφίλ οδήγησης είναι σε θέση να περιγράψουν τη συνολική συμπεριφορά οδήγησης που υιοθετεί κάποιος κατά τη διάρκεια του ταξιδιού του.
- Υπάρχουν δύο κατηγορίες οδηγών σύμφωνα με τη μέση συμπεριφορά του κάθε οδηγού που προκύπτει από τον τρόπο με τον οποίο οδηγούσε σε όλες τις διαδρομές του. Στην πρώτη κατηγορία οι οδηγοί οδηγούν συνήθως με τυπικό τρόπο, ενώ στη δεύτερη κατηγορία οι οδηγοί εκτελούν έναν αριθμό μη ασφαλών ενεργειών οδήγησης ή οδηγούν με επιθετικό τρόπο στην πλειονότητα των ταξιδιών τους.
- Η προσέγγιση Actor-critic από την οικογένεια των αλγορίθμων ενισχυτικής μάθησης μπορεί να αξιοποιηθεί για την εύρεση της καλύτερης δυνατής ενέργειας οδήγησης για κάθε οδηγό με δεδομένο τον τρόπο που οδήγησε στην προηγούμενη διαδρομή του.
- Όταν ένας ελεγκτής παρέχει συστάσεις οδήγησης σε έναν στόλο οχημάτων, το προφίλ επιτάχυνσης ολόκληρου του στόλου εναρμονίζεται σε μια τιμή που είναι αρκετά κοντά στη συμπεριφορά, και συγκεκριμένα στην επιλογή επιτάχυνσης, ενός τυπικού - ασφαλούς οδηγού.
- Η εφαρμογή ενός συστήματος εξατομικευμένων συστάσεων στο οδικό δίκτυο μιας πόλης δεν έχει σημαντικές επιπτώσεις στις συνθήκες κυκλοφορίας.
- Όταν κάθε οδηγός βελτιώνει τη δική του συμπεριφορά, η οδική ασφάλεια στο δίκτυο ενισχύεται. Συγκεκριμένα, οι κρίσιμες εμπλοκές μεταξύ οχημάτων μειώνονται σημαντικά μετά την εφαρμογή του προτεινόμενου συστήματος.
- Η στάθμη των εκπομπών για όλα τα διαφορετικά είδη ατμοσφαιρικών ρύπων μειώνεται, γεγονός που δείχνει ότι η εναρμόνιση των επιταχύνσεων για το σύνολο της κυκλοφορίας μπορεί να έχει σημαντική θετική επιρροή στις περιβαλλοντικές συνθήκες.

Πρέπει να σημειωθεί ότι η επιδείνωση της κυκλοφορίας μπορεί να θεωρηθεί αποδεκτή εάν ληφθεί υπόψη η αντιστάθμιση μέσω των πλεονεκτημάτων από την υιοθέτηση ομαλότερης συμπεριφοράς οδήγησης στην οδική ασφάλεια και τις εκπομπές ρύπων. Για το σκοπό αυτό, οι υπεύθυνοι χάραξης πολιτικής και οι ερευνητές δεν θα πρέπει να παραμελούν τις πραγματικές επιπτώσεις σε όλες τις διαστάσεις του δικτύου όταν σχεδιάζουν στρατηγικές διαχείρισης της κυκλοφορίας και εφαρμόζουν ήπιες και σκληρές πολιτικές (soft and hard policy measures).

Η παρούσα διδακτορική διατριβή συνεισφέρει σημαντικά σε πέντε τομείς:

1. Κάνει χρήση ενός καινοτόμου συνόλου δεδομένων πραγματικής οδήγησης. Ένας σημαντικά μεγάλος όγκος δεδομένων με υψηλή χρονική ανάλυση από πραγματική οδήγηση ήταν διαθέσιμος, εμπλουτισμένος με ποικίλους παράγοντες που περιγράφουν τη συμπεριφορά οδήγησης, το περιβάλλον και άλλα εξωτερικά χαρακτηριστικά για κάθε διαδρομή.
2. Προτείνει ένα μεθοδολογικό πλαίσιο για την εξαγωγή προφίλ οδήγησης απευθείας από τα δεδομένα, τα οποία περιγράφουν όλο το φάσμα της συμπεριφοράς οδήγησης. Για

το σκοπό αυτό, ακολουθείται μια προσέγγιση με βάση τα δεδομένα (data driven approach) για την ομαδοποίηση των κρίσιμων προφίλ οδήγησης που εμφανίζονται κατά τη διάρκεια ενός ταξιδιού, αξιοποιώντας την ομαδοποίηση k-means ως το καταλληλότερο εργαλείο.
3. Αναπτύσσει έναν αλγόριθμο ενισχυτικής μάθησης για την επίλυση ενός πραγματικού προβλήματος, αυτού της υποβοήθησης της οδήγησης. Ένας αλγόριθμος Βαθιάς Ενισχυτικής Μάθησης επιλέχθηκε ως το καταλληλότερο εργαλείο για την εκμάθηση της βέλτιστης πολιτικής και την πρόταση της κατάλληλης ενέργειας που οδηγεί στην καλύτερη δυνατή συμπεριφορά οδήγησης για κάθε μεμονωμένο οδηγό.
4. Προτείνεται μια μεθοδολογία η οποία είναι ικανή να αναγνωρίζει τις ατομικές προτιμήσεις οδήγησης και να παράγει εξατομικευμένες ενέργειες οδήγησης σε κάθε οδηγό. Συγκεκριμένα, υλοποιείται ένα περιεκτικό μεθοδολογικό πλαίσιο το οποίο ενσωματώνει εργαλεία και μεθόδους που πρώτα αναγνωρίζουν τη συμπεριφορά οδήγησης κάθε χρήστη, στη συνέχεια αντιστοιχίζουν κάθε χρήστη στην κατάλληλη έκδοση του μοντέλου RL με βάση τη συνολική συμπεριφορά του και τέλος παράγουν εξατομικευμένες ενέργειες οδήγησης που μετριάζουν την επιθετικότητα και την επικινδυνότητα της οδήγησης.
5. Αξιολογεί τις επιπτώσεις μεγάλης κλίμακας από την εφαρμογή ενός εξατομικευμένου συστήματος συστάσεων οδήγησης σε τρεις τομείς ενδιαφέροντος του δικτύου με τη χρήση συγκεκριμένων δεικτών απόδοσης (KPIs), και συγκεκριμένα στην κυκλοφορία, την ασφάλεια και τις εκπομπές ρύπων. Η αξιολόγηση των επιπτώσεων του προτεινόμενου συστήματος συστάσεων πραγματοποιείται με τη χρήση ενός ρεαλιστικού σεναρίου προσομοίωσης που αφορά το οδικό δίκτυο της Αθήνας και με την εφαρμογή μιας μεθοδολογίας «πριν και μετά» για τη σύγκριση των τιμών των KPIs πριν και μετά την εφαρμογή των συστάσεων οδήγησης.

Περιορισμοί έρευνας, επιπτώσεις και μελλοντική έρευνα

Όπως κάθε άλλη προσέγγιση που βασίζεται σε δεδομένα, έτσι και αυτή η έρευνα στηρίχθηκε σε ορισμένους περιορισμούς όσον αφορά τη δημιουργία και την προσαρμογή του προς έρευνα προβλήματος. Πρώτον, ορισμένοι περιορισμοί προέκυψαν από την ανάγκη να ταιριάξουν τα αποτελέσματα του μοντέλου RL με τους περιορισμούς της προσομοίωσης. Πιο συγκεκριμένα, μία από τις δύο συνιστώσες της συνιστώμενης δράσης, η μέση επιτάχυνση κάθε οδηγού, δεν μπορούσε να εισαχθεί στο μοντέλο μικρο-προσομοίωσης, μιας και αυτό λαμβάνει ως παράμετρο μόνο την ικανότητα επιτάχυνσης των οχημάτων, και επομένως μόνο η μέγιστη επιτάχυνση υιοθετείται εντός της προσομοίωσης. Παρόλα αυτά, λόγω της φύσης του φαινομένου της οδήγησης, όλες οι παράμετροι που περιγράφουν τον τρόπο με τον οποίο ένας οδηγός επιλέγει να οδηγήσει κατά τη διάρκεια μιας διαδρομής είναι άρρηκτα συνδεδεμένες μεταξύ τους και, ως εκ τούτου, η παραμέληση της μέσης επιτάχυνσης δεν αναμενόταν να έχει σημαντική επίδραση στα αποτελέσματα της προσομοίωσης. Πέρα από τον τρόπο με τον οποίο οδηγός ακολουθεί το προπορευόμενο όχημα, ο ίδιος κατά τη διάρκεια των ταξιδιών του λαμβάνει ενέργειες σχετικά με την αλλαγή λωρίδας, την παραχώρηση προτεραιότητας και άλλες αποφάσεις που αφορούν στην αλληλεπίδραση με άλλους χρήστες της οδού. Ωστόσο, στην παρούσα έρευνα η έμφαση δόθηκε ρητά στη συμπεριφορά που σχετίζεται με το προπορευόμενο όχημα, καθώς ο απώτερος στόχος ήταν η δημιουργία ενός συστήματος με επίκεντρο τον χρήστη, το οποίο εξετάζει μόνο τον οδηγό και δεν απαιτεί καμία εξωτερική πληροφορία από το οδικό δίκτυο προκειμένου να εκπαιδευτεί και να εφαρμοστεί. Έτσι, οι προτεινόμενες ενέργειες αναφέρονται στον τρόπο με τον οποίο ο οδηγός κινείται στο δρόμο, δηλαδή στον τρόπο που επιλέγει να πατήσει το πεντάλ επιτάχυνσης, ο οποίος εξαρτάται μόνο από τις προσωπικές προτιμήσεις και αντιλήψεις του οδηγού. Η αγνόηση των πληροφοριών σχετικά με το περιβάλλον μπορεί να θεωρηθεί ως περιορισμός του συστήματος που αναπτύχθηκε, δεδομένου ότι η μετατροπή του σε ένα σύστημα με επίγνωση του περιβάλλοντος θα έδινε άλλες προοπτικές, τόσο στο ίδιο το σύστημα, όσο και στις δυνατότητες χρήσης του ως εργαλείο διαχείρισης της κυκλοφορίας.

Μια προέκταση του παραπάνω περιορισμού είναι το γεγονός ότι, εφόσον το σύστημα αγνοεί την κατάσταση του περιβάλλοντος, δεν μπορεί να λειτουργήσει σε πραγματικό χρόνο. Με άλλα λόγια, η προτεινόμενη μεθοδολογία δεν είναι σε θέση να παράγει συστάσεις σε πραγματικό χρόνο, δηλαδή κατά τη διάρκεια ενός ταξιδιού. Αντ' αυτού, αναπτύσσεται ένα σύστημα που δε λειτουργεί σε πραγματικό χρόνο το οποίο προτείνει αλλαγές στη συμπεριφορά οδήγησης σε μια ακολουθία ταξιδιών για κάθε οδηγό. Η ενσωμάτωση εξωτερικών πληροφοριών στο σύστημα θα επέτρεπε, τουλάχιστον εννοιολογικά, την παροχή συστάσεων οδήγησης σε πραγματικό χρόνο.

Τέλος, ένας άλλος περιορισμός, ο οποίος ισχύει για όλες τις προσεγγίσεις που βασίζονται σε δεδομένα, είναι η γενίκευση και η δυνατότητα μεταφοράς του μοντέλου που αναπτύχθηκε και των αντίστοιχων αποτελεσμάτων. Στις περισσότερες περιπτώσεις δεν είναι σαφές εάν το δείγμα που χρησιμοποιήθηκε για την εκπαίδευση του μοντέλου είναι αντιπροσωπευτικό του συνολικού πληθυσμού και επιπλέον εάν τα χαρακτηριστικά του είναι παρόμοια με εκείνα ενός διαφορετικού πληθυσμού. Στην παρούσα εργασία, για την ανάπτυξη των μοντέλων RL χρησιμοποιείται ένα σύνολο δεδομένων οδήγησης μεγάλης κλίμακας, το οποίο περιλαμβάνει διαδρομές που πραγματοποιούνται από μεγάλο αριθμό οδηγών, ωστόσο δεν μπορεί να ειπωθεί ότι τα αποτελέσματα μπορούν να γενικευτούν και να μεταφερθούν χωρικά σε άλλο οδικό δίκτυο.

Εκτός από τους περιορισμούς που περιγράφονται παραπάνω, τα αποτελέσματα που εξήχθησαν στο πλαίσιο της παρούσας διατριβής μπορούν να έχουν σημαντική επιρροή σε διάφορες πτυχές τόσο της έρευνας (R), της τεχνολογίας (T) όσο και της χάραξης πολιτικής (P). Η μελλοντική έρευνα μπορεί να ωφεληθεί και να εξελιχθεί σημαντικά εξετάζοντας περαιτέρω τα συμπεράσματα που εξάγονται σε σχέση με τα ακόλουθα σημεία:

- (R) Η επιθετικότητα δεν αποτελεί απαραίτητα μια μη ασφαλή συνήθεια οδήγησης και μπορεί να ανιχνευθεί, είτε ως μεμονωμένη συμπεριφορά είτε σε συνδυασμό με άλλες μη ασφαλείς συμπεριφορές.
- (R) Οι αλγόριθμοι ενισχυτικής μάθησης μπορούν να εφαρμοστούν σε προβλήματα του πραγματικού κόσμου και συγκεκριμένα, ο αλγόριθμος DDPG μπορεί να μάθει πώς να

λαμβάνει αποφάσεις όπως ο άνθρωπος σε πολύπλοκα και υψηλών διαστάσεων περιβάλλοντα.

- (R & T) Τα αναγνωρισμένα προφίλ οδήγησης μπορούν να παρέχουν σημαντική γνώση για την ανάπτυξη μοντέλων αυτόνομης οδήγησης που θα προσομοιάζει την ανθρώπινη οδήγηση.

Όσον αφορά στην τεχνολογία, το προτεινόμενο σύστημα συστάσεων μπορεί να ενσωματωθεί σε ήδη αναπτυγμένα λογισμικά, όπως εφαρμογές τηλεματικής και συστήματα ADAS μετατρέποντάς τα σε πιο φιλικά προς το χρήστη και να προσφέρει έναν πιο εξατομικευμένο τρόπο υποστήριξης της λήψης αποφάσεων κατά την οδήγηση. Επιπλέον, οι υπεύθυνοι χάραξης πολιτικής θα μπορούσαν να επωφεληθούν από τα αποτελέσματα της παρούσας διατριβής για τον επανασχεδιασμό των μέτρων ήπιας πολιτικής και τον επαναπροσδιορισμό του ρόλου των οδηγών στις τρέχουσες στρατηγικές διαχείρισης της κυκλοφορίας, δεδομένου ότι στην παρούσα εργασία αποδείχθηκε ότι η βελτίωση της συμπεριφοράς οδήγησης σε ατομικό επίπεδο μπορεί να έχει σημαντική επίδραση στην οδική ασφάλεια και τις εκπομπές ρύπων, αλλά όχι αξιοσημείωτες επιπτώσεις στις συνθήκες κυκλοφορίας.

Τέλος, μπορεί να γίνει κατανοητό ότι τα ευρήματα αυτής της εργασίας προσφέρουν σημαντικές κατευθύνσεις για τη μελλοντική έρευνα. Παρόλο που η παρούσα έρευνα συνεισφέρει σημαντικά στην ανάλυση της συμπεριφοράς οδήγησης, υπάρχουν ακόμη περιθώρια στη διερεύνηση της δυναμικής της και, συνεπώς, θα πρέπει να διεξαχθεί περαιτέρω έρευνα που θα περιλαμβάνει εμπλουτισμένα σύνολα δεδομένων οδήγησης με πρόσθετες συμπεριφορές κατά την οδήγηση (π.χ. εργασίες που προκαλούν απόσπαση της προσοχής εκτός από τη χρήση κινητού). Επιπλέον, η εξειδικευμένη διερεύνηση της δυναμικής εξέλιξης της συμπεριφοράς οδήγησης είναι επίσης πολύ σημαντική για να δοθούν απαντήσεις στο ερώτημα αν και πόσο γρήγορα μεταβάλλονται τα προφίλ οδήγησης με την πάροδο του χρόνου για κάθε οδηγό. Μια άλλη κατεύθυνση μελλοντικής έρευνας αφορά στο σύστημα συστάσεων, και αφορά τον τρόπο με τον οποίο οι παραγόμενες συστάσεις θα πρέπει να μεταβιβάζονται στον οδηγό ώστε να γίνονται κατανοητές και στη συνέχεια να γίνονται αποδεκτές από αυτόν. Επιπλέον, ο προσδιορισμός των απαιτούμενων προδιαγραφών που θα επιτρέψουν τη λειτουργία του συστήματος σε πραγματικό χρόνο θα μπορούσε επίσης να αποτελέσει μέρος της μελλοντικής έρευνας. Προς αυτή την κατεύθυνση, ο σημαντικότερος μελλοντικός ερευνητικός στόχος θα ήταν η τροποποίηση του προτεινόμενου συστήματος κατά τρόπο ώστε να αποκτήσει επίγνωση του περιβάλλοντος, δηλαδή το σύστημα να μπορεί να αλληλεπιδρά με το περιβάλλον στο οποίο ο πράκτορας λαμβάνει αποφάσεις, και να έχει πλήρη εικόνα της δυναμικής και των μεταβολών του. Με αυτόν τον τρόπο, το προτεινόμενο σύστημα θα μπορούσε να εφαρμοστεί σε πραγματικό χρόνο και επιπλέον θα μπορούσε να λειτουργήσει ως εργαλείο διαχείρισης της κυκλοφορίας, το οποίο χρησιμοποιεί τη συμπεριφορά των οδηγών ως βασική δύναμη ενίσχυσης της αποδοτικότητας της κυκλοφορίας.

Extended Summary

The urban transportation landscape is facing many challenges due to the introduction of a variety of mobility solutions for travelers which together with innovations of Information and communication Technologies (ICT) subvert traffic management policies. Policy makers have to reconsider the applied traffic management measures in a way that automation and cooperation requirements of today's services are taken into account together with the always increasing needs for green and sustainable mobility solutions. Nevertheless, even in this ever-changing transportation system, drivers remain the protagonists. Therefore, the understanding of decision-making while driving, as well as the investigation of driving habits adopted by drivers remains an active field of research for more than a decade.

Abnormal driving has been linked with increased crash risk and, thus, the improvement of driving behavior is considered critical for improving road safety. In addition, previous research has implied that the improvement of individual driving behavior may also result in an improvement of traffic conditions. Nevertheless, no evidence has been provided to support this statement, and the consequences of adjusting individual driving behavior on a network-level still remain unclear.

Within this context, the work contained in this dissertation is motivated by two main driving forces: i) the need to develop a driving recommendation system that treats each driver as an individual and proposes actions that meet his/her own driving preferences and, ii) the need to explore the actual impact of applying a personalized recommendation system on the road network.

Main objectives and research questions

The main objective of this dissertation is to design a personalized driving recommendation system which is based on deep reinforcement learning algorithms and aims at enhancing driving safety through the mitigation of aggressiveness and other unsafe driving habits. Subsequently, the impact of controlling individual driving behavior is assessed with regards to network performance and road safety, as well as the levels of harmful emissions by properly adjusting parameters of traffic models in a city-wide scenario setting using microsimulation. The above-described overarching goal of this dissertation can be divided in three major objectives as described below:

1. Exploit smartphone sensed data to understand driving behavior
2. Develop a traffic theory compatible personalized recommendation framework for improving driving behavior
3. Assess the impact of the recommendation system in traffic, safety and emissions

The concept of driving behavior analysis is not new, and, thus, a thorough review of the literature was conducted, at first, with the aim to identify research gaps and highlight the

challenges and caveats that arise when smartphone crowd-sensed data are exploited for this purpose. The review of the literature resulted in the formation of the following 7 research questions.

Question 1 (Q1): Which are the main driving profiles that cover the wide range of driving behavior and how can they be identified by exploiting smartphone data?
Question 2 (Q2): Is it possible to classify the overall driving behavior of drivers into groups that share common driving characteristics, and, if so, to what extent could it be classified?
Question 3 (Q3): Could Artificial Intelligence techniques be exploited within the framework of a driving recommendation system and ensure the requires degree of personalization of the produced recommended actions?
Question 4 (Q4): Which is the most appropriate Reinforcement Learning algorithm for supporting human decision making?
Question 5 (Q5): Is there a link between raising self-awareness and improving conditions of the entire network? To what extent could the improvement of individual behavior affect traffic conditions?
Question 6 (Q6): What kind of impact does the controlling of individual driving behavior have on driving and road safety?
Question 7 (Q7): How are emissions affected by the controlling of individual driving behavior? Is there a significant change on environmental conditions when drivers improve their behavior?

Methodological approach

The recommendation system proposed within this dissertation is basically a decision support system for drivers that aims at mitigating aggressiveness and riskiness. Driving is a complex task since it requires from the driver to take both strategic and dynamic decisions as well as adapt their behavior to emerging conditions of the network. Contrary to the already developed ADAS, the system here has the following three state-of-the-art characteristics:

1. It is personalized, which means that it recommends the best driving actions to each individual taking into account their specific requirements and driving preferences.

2. It is self-aware, which means that the system takes into account previous behavior of each individual driver in order to propose the most suitable driving recommendations.

3. It is autonomous, meaning that it does not require any external input from the network or the traffic. Driving recommendations aim to improve individual driving behavior on its core, namely acceleration and deceleration decisions.

The development of the recommendation system is based on a Reinforcement Learning algorithm which is capable of producing the optimal behavior alteration for each driver given the way they have drove over their last trip.

In order to answer the research questions and achieve the overarching goal of the dissertation, an inclusive methodological framework is proposed which is based on a mixture of unsupervised learning and Deep Reinforcement Learning algorithms as depicted in Figure I.

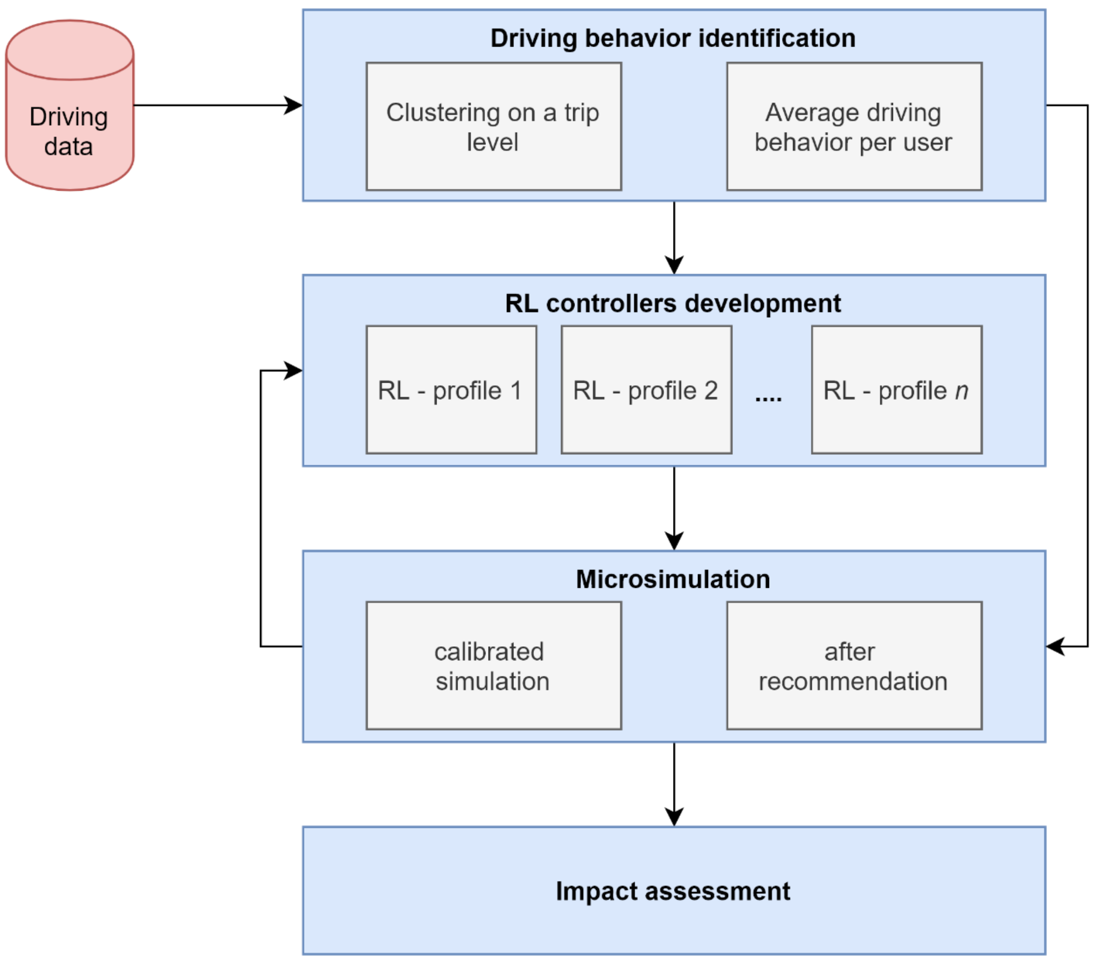


Figure I. Overview of the methodological framework

Starting from raw measurements of GPS location, acceleration and speed, as provided by a telematics application established on smartphone devices, driving features are defined that describe short-term and long-term driving behavior. Following, these features are utilized in an unsupervised learning framework to identify driving profiles that can be used to describe each driver's overall driving behavior (Q1, Q2). Driving behavior is defined at:

- a trip level, which corresponds to the way the driver performed a specific trip, and
- a user level, which corresponds to the overall driving behavior of a specific driver in all of his trips (driving footprint).

A two-level k-means clustering algorithm is implemented, in a selection of driving features, in order to distinguish aggressive from non-aggressive trips within the first level, and then further distinguish between risky and distracted driving at the second level of clustering. After this procedure, each trip was assigned to a specific driving profile (Q1), and then, using statistical measurements, the overall driving behavior of each driver is identified (Q2).

Once driving behavior per trip is identified, and all drivers were separated into groups based on their overall driving behavior, the driving recommendation framework is designed and the appropriate algorithms are developed, using state-of-the-art Reinforcement Learning models (Q3). The aim of the Reinforcement Learning algorithm is to learn the optimal policy and suggest the appropriate action that leads to the best possible behavior. Specifically, when dealing with driving behavior recommendations, every action refers to an adjustment of the vehicle's kinematic characteristics including the adjustment of the vehicle's speed and acceleration, which span within a continuous range of values. To this end, the RL algorithm developed in this work should retain one extra property, the ability to handle continuous state and action spaces (Q4). The RL agents follow an actor-critic approach based on the Deep Deterministic Policy Gradient algorithm and are both implemented as deep artificial neural networks, the hyperparameters and the structure of which emerge after an exhaustive grid search. The algorithms are trained using sequences of driving trips of the same driver as input, while the output of each RL controller is the optimal alteration in the acceleration of each driver, given the way they drove in their previous trip.

The structure of the system is such that there is a full mapping to microscopic standardization and traffic flow control logic. In correspondence with widely used car following models, the proposed algorithm acts as an estimation and prediction function of the acceleration at which the vehicle should move.

Finally, the impact of improving individual driving behavior is assessed through a comparative before-after microsimulation analysis, with respect to road safety, traffic and the environment (Q5, Q6, Q7). Using the road network of Athens, Greece, a microsimulation scenario for the morning rush hour demand, was set. For the initial conditions of the network the vehicles move according to the characteristics governing each of the driving behaviors detected in the first step of the methodological framework, while the traffic composition is based on the actual distribution of trips over the driving profiles. In this way, driving diversity is ensured between the vehicles and the traffic conditions in the network are simulated as realistically as possible.

The data

For the purpose of the specific research, data were collected through an innovative smartphone application developed by Oseven Telematics. The naturalistic driving database included 153,953 trips made from 696 unique drivers from December 2017 to August 2019. The trips were performed all around Greece, nevertheless the majority of them were conducted within the Region of Attica. For each trip, a variety of variables are available which include statistical measurements of acceleration and deceleration during a trip, speeding measurements that describe smoothly and with speed excess driving, as well as mobile usage indicators that describe how cautious the driver is. Table I presents the driving parameters used in the specific research.

Table I. Driving parameters per trip

Variable	Description	Unit
harsh_acc_per_min	Average number of harsh accelerations performed per minute	events/min
acc_avg	Average acceleration	m/s^2
acc_std	Standard deviation of acceleration	m/s^2
acc_q90	90% percentile of acceleration	m/s^2
acc_max	Maximum acceleration	m/s^2
harsh_brk_per_min	Average number of harsh decelerations performed per minute	events/min
dec_avg	Average deceleration	m/s^2
dec_std	Standard deviation of deceleration	m/s^2
dec_q90	90% percentile of deceleration	m/s^2
dec_max	Maximum deceleration	m/s^2
speed_max	Maximum speed	km/h
mbu	Percentage of driving with mobile usage	%
speeding_percentage	Percentage of driving with speed over the speed limit	%

All data provided by Oseven are in a fully anonymized format. The main characteristics of the sample used in the specific research are presented in Table II.

Table II. Main characteristics of the sample used in this research

	Total	Safe	Unsafe
Number of trips	153,953	66,566	87,387
Number of drivers	696	197	499
Average number of trips per driver		221	
Minimum number of trips per driver		16	
Average km travelled per driver		2,510 km	

Driving behavior analysis

In order to achieve the first objective of this dissertation, which is to exploit smartphone sensed data to understand driving behavior, a k-means clustering algorithm is implemented in to distinct levels.

For the first level of clustering, the number of clusters is set to k=2 and clustering is implemented on Euclidean distance matrix. Two of the variables that are used for the above procedure describe the number of harsh alterations of the longitudinal position of the vehicle (acceleration and deceleration), while the rest of them are essentially indices of the average acceleration and deceleration of the trip. The results of this first implementation of the k-means clustering are presented in Table III.

Table III. 1st level clustering results

	Harsh acceleration per min	Harsh brake per min	Average acceleration	Standard deviation of acceleration	Maximum acceleration	Average deceleration	Standard deviation of deceleration	Maximum deceleration	Number of trips
Aggressive trips	0.150	0.2081	1.748	1.525	3.847	-1.968	1.843	-4.547	71263
Non-aggressive trips	0.028	0.051	1.137	1.052	2.503	-1.282	1.286	-2.926	82690

Based on the clusters' centers, the trips can be distinguished between aggressive and non–aggressive driving, since trips belonging to the first cluster are featured by aggressive driving characteristics, such as great acceleration and deceleration metrics and significantly higher rates of harsh events per minute of driving.

The second level of k-means clustering was applied separately to the two groups that emerged from the first level of clustering using two driving parameters: the percentage of driving with mobile usage and the percentage of driving with speed over the speed limit. Results of this second level of clustering are presented in the table below (Table IV).

Table IV. 2nd level clustering results

	Percentage of mobile usage	Percentage of driving with speed over the speed limit	Number of trips
		Aggressive trips	
Distracted	0.511	0.062	4505 (2.9%)
Aggressive	0.019	0.032	54394 (35.3%)
Risky	0.023	0.269	12364 (8%)
		Non-aggressive trips	
Risky	0.021	0.306	12494 (8.1%)
Moderate	0.014	0.029	66566 (43.2%)
Distracted	0.514	0.057	3630(2.4%)

The resulting clusters seem to reveal richer driving profiles: distracted driving is recognized by higher values of the percentage of mobile usage while driving, while risky driving is identified through higher values of percentage of driving with speed over the speed limit. The two remaining clusters which have the lower values in both measures are annotated as "aggressive" and "moderate" for the aggressive and non-aggressive trips subsets respectively.

In order to separate drivers into groups with the same driving preferences, an average driving profile of each individual was identified by applying a simple rule. All four driving profiles indicating an unsafe driving behavior (Risky, Distracted, Aggressive-risky, Aggressive-distracted) were grouped as the worst class (3), aggressive trip profiles constitute the second class (2), while trips with typical characteristics belong to the first class (1), as shown in Figure II. For each individual driver, an average from all their trips is estimated and drivers are separated into two main groups based on their average behavior, as follows:

- Moderate/typical drivers: *trip average ≤ 1.5*
- Reckless drivers: *trip average > 1.5*

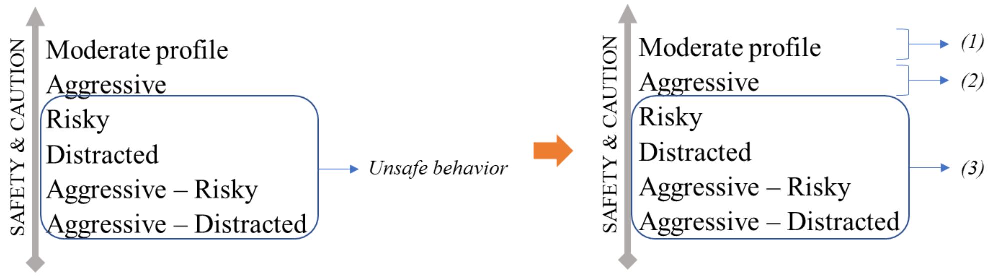


Figure II. Trip profile grouping for drivers' average driving profile estimation

For each individual driver, an average of the annotations from all their trips is estimated, where trip average less than 1.5 implies a moderate/typical driver and trip average greater than 1.5 refers to reckless drivers. Based on some statistical analysis, trip average less than 1.5 indicates that at least 60% of the trips performed by a driver are characterized by "moderate" driving behavior. In order for the developed controller to be as adaptive as possible to each individual's behavior, the proposed framework should be very strict when characterizing a driver as "typical/moderate" in order to avoid suggesting changes in behavior that the driver himself is impossible to follow as they will be far from his own average behavior.

RL: concept, principles and model development

In order to develop the Self-Aware Driving Recommendation Assistant (SADRA), a structured procedure is followed. First, the total trip database is divided into two, based on the average driving profile of each driver. In particular, the first database includes the trips of all drivers belonging to the "typical-safe" drivers, while the second includes all the trips of drivers with unsafe average driving behavior. For the sake of brevity, from this point on, the RL controller that corresponds to the "typical" drivers is referred to as SADRA – I, while the corresponding controller for the reckless drivers is referred to as SADRA – II respectively.

Every RL agent consists of three main components: *states (s)*, *actions (a)* and *rewards (r)*. In each timestep the agent observes the current state of the environment and takes the appropriate action from the set of the possible actions. Then, the agent receives a reward which measures the success or failure of the agent's actions for the given state.

In this study, the environment states are defined through a five-dimensional vector that describes how a driver drove during their trip and includes trip's average acceleration (a_{avg}), 90% percentile of acceleration (a_{90}), average deceleration (d_{avg}), 90% percentile of deceleration (d_{90}) and percentage of driving with speed over the speed limit (*speeding*):

$$\boldsymbol{s} = \{a_{avg}, a_{90}, d_{avg}, d_{90}, speeding\}$$

Our recommendation system is not context-aware which means that its ultimate goal is to improve individual's personal driving style independently from the road setting they are driving in (type of road, traffic conditions, etc.). The selection of the appropriate speed is not independent from the road geometry and road traffic, as well as deceleration decisions are not always independent from the leading vehicle's behavior and traffic signals. Therefore, the only parameter that purely describes one's driving style is the acceleration, as it is only dependent on the driver's perception and preference between smoothly or harshly accelerating. Indeed, in recent literature, a driver's driving style is usually defined by their acceleration profile. To this end, actions that the system produces and are proposed to the driver belong to a continuous action space which is defined by a two-dimensional vector including a change in average acceleration and in the 90% percentile of acceleration, which define the usual/preferred acceleration for the entire trip in regular situations and the value that should not be exceeded, e.g., when performing overtaking maneuvers, except from cases of emergency:

$$\boldsymbol{a} = \{da_{avg}, da_{90}\}$$

For the sake of simplicity from hereon, the 90% quartile of the acceleration may be equally referred to as "maximum acceleration".

A key component of the RL agent is the reward function. The aim of the reward function is twofold; to evaluate the current state and the transition between states. In other words, the driving behavior at each trip, as well as the change in driving behavior between successive trips of the same user are evaluated. For this purpose, a custom driving evaluation function had to be constructed first. The score of each trip was estimated by the distance of this specific trip from the center of the moderate profile (the center of the cluster), in order to quantify how far each individual's behavior is from the typical (moderate) behavior. For the purpose of this analysis, the Mahalanobis distance is used to estimate the distance between each trip and the moderate profile.

Trip evaluation is performed on the basis of the following formula:

$$trip\ score_i = e^{-driving\ profile_i * \frac{Ɱ_{(i,moderate\ profile)}}{Q_{75}(Ɱ)}}$$

where i is an individual trip and Ɱ is the Mahalanobis distance. Here, the 3rd quartile of the Mahalanobis distance is used instead of the maximum value in order for the score function to be stricter with drivers whose behavior excludes more than 75% of the typical (moderate) behavior.

The reward function for a driver moving from one trip to the next one was established based on the following formula:

$$\boldsymbol{r} = \ trip\ score_{i+1} \left(1 + \frac{trip\ score_{\mathrm{i+1}} - trip\ score_{i}}{100}\right)$$

Once the main components for the development of the RL controllers were estimated, the data were organized in the following format:

$$(state, action, reward, next\ state)$$

For every unique driver in the dataset, their trips were sorted in an ascending order according to each trips starting date. The training samples were tuples of sequential trips of a specific driver along with the corresponding action and reward of the transition from the first trip to the succeeding one. It should be noted that for every distinct driver in the dataset, their first trip was used only as "*state*" while their last trip of was used only as "*next state*". Following this data preparation procedure, 33,440 unique data samples were constructed for training SADRA I and 119,817 unique data samples were used for the training process of SADRA II.

The RL controllers are developed based on the Deep Deterministic Policy Gradient (DDPG) algorithm which implements an actor-critic approach to learn a policy and produce the optimal actions. Thus, for each controller two neural networks are developed; representing the actor and the critic respectively. The actor (μ) and critic (Q) networks for both the safe and unsafe drivers' subsets were trained following the procedure of Algorithm below.

DDPG Algorithm implementation

Initialize critic $Q(s,a|\theta^{Q})$ and actor $\mu(s|\theta^{\mu})$ networks using rewards as Q-values
Set the above as initial target networks (Q' and μ')
Split the sample into M minibatches
for minibatch=1, M **do**
  Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$
  Update critic by minimizing the loss: $L = \frac{1}{N}\sum_i \left(y_i - Q(s_i, a_i|\theta^{Q})\right)^2$
  Update the actor policy using the sampled policy gradient:

$$\nabla_{\theta_\mu} J \approx \frac{1}{N}\sum_t [\nabla_a Q(s,a|\theta^{Q})|_{s=s_t, a=\mu(s_t)}\ \nabla_{\theta_\mu}\mu(s|\theta^{\mu})|_{s=s_t}]$$

  Update the target networks:

$$\theta^{Q'} \leftarrow \tau\theta^{Q} + (1-\tau)\theta^{Q'}$$
$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'}$$

end for

An exhaustive grid search was performed in order to conclude to the final architecture of the two networks. Specifically, all possible combinations of the networks' structures and parameterization, within a range of reasonable values, have been examined and compared in order to detect the optimal one. The parameters that were taken into consideration are: number of hidden layers, number of neurons and activation of each layer, optimization algorithm and learning rate, batch size and number of training epochs, as shown in Table V.

Table V. Hyperparameters of the Critic and Actor networks for both SADRA I and II

Hyperparameters	Critic network	Actor network
	SADRA I – Safe drivers	
Number of hidden layers	6	3
Number of neurons per layer	(64,32,16,16,32,64,1)	(128,64,32,2)
Epochs	200(initial network:110)	200(initial network:110)
Batch size	150(initial network:150)	150(initial network:150)
Activation	ReLU	ReLU
Optimizer	Adam	Adam
Learning rate	0.001	0.0001
	SADRA II – Unsafe drivers	
Number of hidden layers	6	3
Number of neurons per layer	(32,16,8,8,16,32,1)	(128,64,32,2)
Epochs	170(initial network:170)	210(initial network:210)
Batch size	250(initial network:250)	250(initial network:100)
Activation	ReLU	ReLU
Optimizer	Adam	Adam
Learning rate	0.0001	0.0001

Simulation setting

The quantification of the impact of adopting driving recommendations by all drivers on traffic, road safety and emissions was performed under a network-level microscopic simulation scenario. The SUMO simulation software is used and its default car-following model, Krauss model, which is a microscopic, space-continuous model based on the safe speed; the driver of the following car adopts a safe speed which allows them to adapt to the deceleration of the leading vehicle.

The case study for the simulation experiments is the inner-ring network of Athens, Greece. The network consists of 1,293 nodes/intersections and 2,572 edges. The total length of the network is 348 kilometers. The calibration of the network led to the definition of **86,054 vehicles**, achieving a total of 1,393,634 counts (97.47% of the total counts extracted from the Aimsun simulator) and a GEH value below 5 (GEH < 5) for 95.26%.

Two distinct scenarios were designed both corresponding to the demand of the Athens' Road network during the morning peak hour (8:00 – 9:00 AM). First, the initial conditions of the network are simulated in order to estimate the performance of traffic when vehicles move around, based on the characteristics that govern the six identified driving profiles. In order to ensure the robustness of the results, simulation was performed in 10 replications with ten different seed numbers. Stochasticity is an important aspect of reproducing reality in a simulation scenario, since it adds randomness over the distributions of difference aspects of the simulation (e.g., route distributions, vehicle type distributions). Subsequently, driving recommendations were produced offline for every served vehicle based on the way each vehicle performed their trip. The recommendations were produced from the corresponding RL controllers using as input the state of the trip (average acceleration, 90% percentile of acceleration, average deceleration, 90% percentile of deceleration, speeding percentage) and as output the optimal alteration of the maximum acceleration. It should be highlighted here that although the developed RL controllers produce a two-dimensional vector that includes alterations on both the average and the maximum acceleration, only the maximum acceleration was exploited during the simulation runs, since the Krauss model takes into account only the maximum values of acceleration and deceleration.

Finally, a second simulation run was performed, where previously served vehicles follow the proposed recommendations, namely an alternation of their maximum acceleration, while the rest of the traffic follows the distribution among the six driving profiles.

The behavior that implies each driving profile was simulated through the adjustment of the car-following model. The car-following model can be parametrized by a number of parameters: the maximum acceleration of the vehicle (*accel*), the maximum deceleration of the vehicle (*decel*), the maximum velocity of the vehicle (*maxSpeed*), the maximal physically possible deceleration for the vehicle (*emergencyDecel*) and the vehicles' expected multiplicator for lane speed limits (*speedFactor*). At first, the current (initial) state of the road traffic is simulated in SUMO using the six defined driving profiles, whose parameters were introduced to the Krauss model of different vehicle types, as shown in Table VI.

Table VI. Car-following model parameters for each vehicle type

Vehicle types (trip profiles)	**Car-Following Model Parameters**				
	accel (m/s^2)	**decel** (m/s^2)	**emergencyDecel** (m/s^2)	**maxSpeed** (km/h)	**speedFactor** (mean, min, max)
Moderate	2.519	-2.942	-5.909	64.51	(0.029, 0, 0.168)
Aggressive	3.817	-4.483	-18.083	66.93	(0.033, 0, 0.151)
Risky	2.392	-2.824	-5.328	100.28	(0.306, 0.1627, 0.96)
Distracted	2.601	-2.990	-5.112	67.38	(0.057, 0, 0.631)
Aggressive-risky	3.944	-4.825	-25.884	100.8	(0.269, 0.147, 0.907)
Aggressive-distracted	3.939	-4.553	-10.845	71.99	(0.062, 0, 0.744)

For the initial state of the network, the six distinct vehicle types were created in a route file, with the corresponding car-following model's parametrization. The route of each vehicle was also identified in the route file, as it was estimated from the path assignment of Aimsun. In one hour of simulation for the morning peak, about 58% of the total demand was inserted in the network and 28% of the vehicles completed their journey within this time.

Subsequently, for each vehicle that reached their destination the following parameters were estimated for each trip:

- average acceleration
- 90% percentile of acceleration
- average deceleration
- 90% percentile of deceleration
- speeding percentage

These driving characteristics were used as input to the RL controllers which recommend the optimal action for each trip. For the second run of simulation, the exact same vehicles were used, which follow the exact same routes on the same road network, in order to estimate the impact of the recommendation. The proposed actions of each vehicle were introduced as a modification of the car-following model's parameter in the route file. The adoption of this approach enabled hands-on implementation of the recommendation process with direct control over the outcomes.

In this case as well, 10 replications with the same seed values as before, were performed to ensure the robustness of the results. Findings revealed that in one hour of simulation 57% of the demand was served on average, while the corresponding percentage of served vehicles was reduced by 1% compared to the initial conditions.

Impact assessment

Impact assessment of the proposed system is performed using microsimulation and by following a before-after approach. Specifically, for both simulation cycles the Key Performance Indicators of traffic, safety and environmental conditions were estimated, and comparatively assessed so that to quantify the overall impact of adopting personalized driving recommendations which improve each individual's driving behavior. The KPIs used in the analysis for each network's aspect are presented in Table VII.

Table VII. Key Performance Indicators for each network's aspect

Traffic	Safety	Environment
Served demand	Total conflicts	Cumulative amount of emissions (CO_2, CO, PMx, NOx)
MFDs	Total rear-end conflicts	Emissions per vehicle
Travel times	Conflicts per vehicle	

The estimation of traffic-related KPIs was dependent on the outputs of the simulation, which included the number of inserted and served vehicles, as well as edge-based information regarding the three fundamental elements of traffic flow theory (flow, speed and density). Instead of using aggregated measures of the fundamental variables, the Macroscopic Fundamental Diagrams (MFDs) were constructed and significant outcomes were drawn regarding the differences in the performance of the network before and after the application of the recommendation system. The estimation of the harmful air pollutants is based on the emissions' model already integrated into SUMO, the PHEMlight model. PHEMlight is a simplified version of PHEM (Passenger car and Heavy-duty Emission Model), a complete vehicle emissions model developed in Europe since 1999. PHEM is based on extensive emission measurements on vehicles such as passenger cars, light duty vehicles and urban buses. The approximation of the conflicts that constitute an indicator for road safety is based on the SSAM tool, which computes a number of surrogate measures of safety for each conflict (crossings, rear-ends, lane changes) that is identified in the trajectory data and then computes summaries (mean, max, etc.) of each surrogate measure.

Results: Driving recommendations

The two versions of the trained DDPG algorithm were used to produce driving recommendations with respect to two categories of drivers; typical drivers who exhibit a moderate average behavior (SADRA I) and unsafe drivers who interchange their behavior among various unsafe driving habits (SADRA II). The recommendations are in the form of driving alterations that refer to the optimal driving actions that the specific driver can adopt in order to improve their driving based on their current behavior.

A comparison between the outputs of the two controllers revealed that both of them are trained to generate recommendations that move drivers closer to the average safe behavior of a typical driver, which has an average acceleration equal to 1.137 m/s^2 and a maximum acceleration equal to 2.503 m/s^2. Based on the indicative samples of the table below (Table VIII), the mean recommended average acceleration was estimated 1.145 m/s^2, while the mean value of the proposed maximum accelerations was 2.507 m/s^2 respectively. It can therefore be concluded that a universal application of the proposed recommendation system would lead to the harmonization of the acceleration profiles for the entire fleet of vehicles.

Table VIII. Example of RL input and output and the produced recommendations

Profiling and Scoring		RL input					RL output		Recommendation	
Driving profile	**Score**	**Acceleration average**	**Deceleration average**	**Acceleration Q90**	**Deceleration Q90**	**Speeding percentage**	**action_1 (max)**	**action_2 (avg)**	**new Acceleration Q90**	**new Acceleration average**
Aggressive-distracted	9	1.512	-1.58	3.600	-3.72	0.0090	↓ -0.99	↓ -0.31	2.610	1.197
Aggressive-distracted	2	1.888	-1.86	3.336	-3.95	0.0660	↓ -0.29	↓ -0.51	3.046	1.377
Aggressive-risky	6	1.779	-1.63	3.432	-3.82	0.3235	↓ -0.65	↓ -0.49	2.778	1.288
Aggressive-risky	2	1.574	-1.73	3.480	-4.20	0.5693	↓ -0.96	↓ -0.40	2.519	1.171
Aggressive	2	1.905	-1.68	3.600	-3.18	0.0000	↓ -0.77	↓ -0.63	2.834	1.275
Aggressive	17	1.948	-2.57	3.768	-5.69	0.0822	↓ -1.10	↓ -0.66	2.669	1.289
Aggressive	10	1.774	-2.17	4.368	-4.08	0.0000	↓ -1.69	↓ -0.54	2.676	1.230
Distracted	2	1.405	-1.33	4.104	-2.88	0.0133	↓ -1.38	↓ -0.21	2.727	1.193
Distracted	24	1.504	-1.52	3.324	-3.36	0.0000	↓ -0.64	↓ -0.27	2.679	1.236
Distracted	35	1.054	-1.08	2.489	-2.67	0.0000	↑ 0.06	↑ 0.11	2.548	1.166
Moderate	73	1.386	-1.59	2.760	-3.28	0.0226	↓ -0.46	↓ -0.33	2.299	1.055
Moderate	58	0.592	-0.92	1.560	-2.81	0.0294	↑ 0.56	↑ 0.39	2.123	0.980
Moderate	86	1.366	-1.43	2.880	-3.36	0.0000	↓ -0.45	↓ -0.26	2.426	1.106
Moderate	90	0.826	-1.03	1.984	-2.43	0.0000	↑ 0.26	↑ 0.18	2.240	1.006
Moderate	67	0.967	-1.43	2.040	-3.67	0.0674	↑ 0.05	↓ -0.01	2.091	0.954
Moderate	78	1.047	-1.23	2.070	-2.33	0.0000	↑ 0.21	↓ -0.02	2.284	1.030
Moderate	77	0.666	-0.69	1.440	-1.68	0.0367	↑ 0.76	↑ 0.32	2.195	0.989
Moderate	94	1.124	-1.24	2.568	-2.64	0.0000	↓ -0.06	↓ -0.01	2.511	1.118
Moderate	58	0.871	-1.24	1.920	-3.56	0.1667	↑ 0.27	↑ 0.13	2.187	1.003
Risky	52	1.367	-1.36	3.093	-3.06	0.2392	↓ -0.49	↓ -0.17	2.602	1.196
Risky	33	1.098	-1.43	2.280	-3.24	0.2857	↑ 0.32	↑ 0.09	2.600	1.192

Figure III provides some indicative examples of the recommendations produced by the two controllers given the same input (first state). Findings revealed that although the recommendations of the controller concerning unsafe drivers (SADRA II) lead to significantly lower average accelerations for the next trip (next state) compared to the previous trip (initial state), they maintain a significant distance upwards for the respective recommendations produced from the typical drivers' RL controller (SADRA I). Nevertheless, it should be noticed that both the controllers lead to a smoother acceleration profile for the entire traffic.

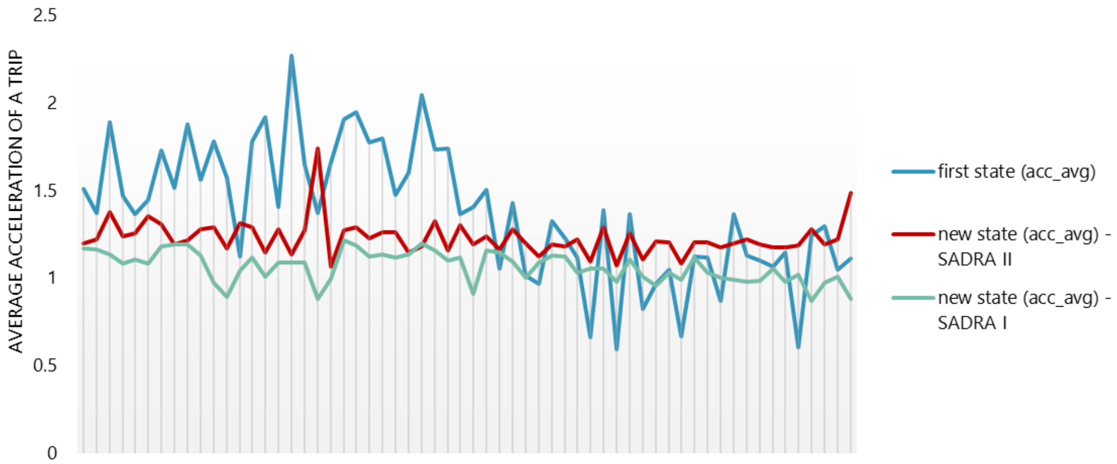


Figure III. Comparison of the new state's average acceleration as it emerged from the Typical and the Unsafe RL controllers.

Results: Simulation and Impact

The quantification of the impact of applying the proposed recommendation system and in consequence, of the adoption of an improved driving behavior by all drivers is of great importance both for researchers as well as practitioners and can lead to significant findings regarding the usefulness of improving individual driving behavior. The assessment of the recommendation system is performed by utilizing specific Key Performance Indicators that correspond to three areas of interest: traffic, safety and emissions. Each of the simulation rounds was done in 10 replications to enhance the validity and robustness of the results. In total, the trained SADRA I controller was used to produce recommendations for 43% of the vehicles, while the rest of the vehicles followed the recommendations produced by SADRA II.

All replications of the same simulation round present mutual results which are slightly reduced after the application of the recommendation system. On average, 2.9% less vehicles were served based on the results of the second round of the simulation. However, results of the statistical hypothesis test *t-test* indicated that there are no significant differences between the means of the served vehicles before and after the recommendations in 95% confidence interval. A greater investigation of the traffic flow properties together with aggregated metrics of driving behavior was conducted to further quantify the impact on the other dimensions of the road network as well. The application of the personalized recommendation system had a substantial impact on the maximum acceleration of the vehicles, as shown in Figure IV. When all vehicles followed the suggestions generated by the two RL controllers, the mean value of the maximum acceleration was somewhat increased from 2.83 m/s^2 to 2.96 m/s^2, mostly because the majority of the vehicles who adopted a very small maximum acceleration, which was far lower from the corresponding acceleration of the "moderate/typical" behavior, they were suggested to slightly increase their acceleration. However, the condensation of the interquartile range is evident after the recommendations, which indicates the harmonization of the acceleration profiles of all vehicles in the simulation. Finally, the maximum value of the observed maximum accelerations remained at the same level of 3.94 m/s^2 after the application of the proposed system.

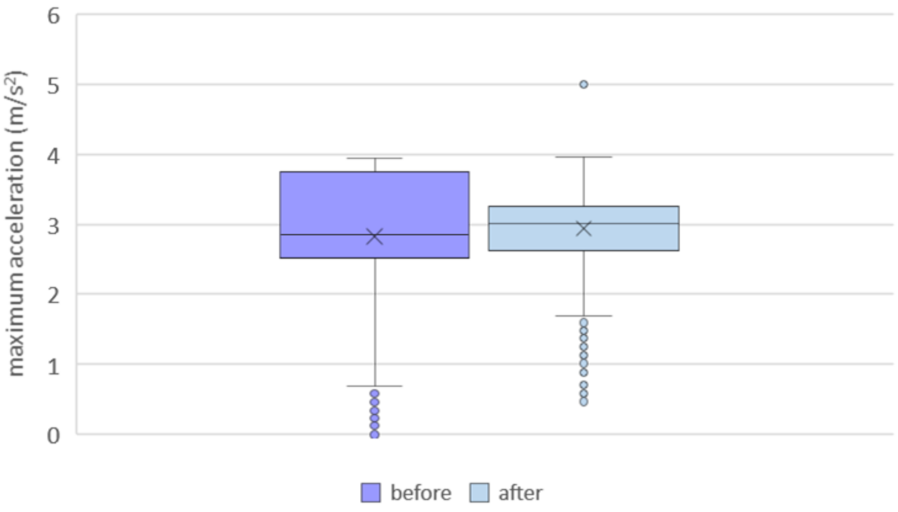


Figure IV. Boxplot of maximum acceleration before and after recommendation

The differences observed in the magnitude of the average speed are minimal, since in both situations the vehicles adopt an average speed of around 25 km/h, while the maximum average speed that is observed is approximately 55 km/h.

Alterations on the speed of vehicles resulted on changes of the rest traffic flows properties, namely flow and density. Microscopic fundamental diagrams were calculated to provide a thorough graphical representation of these variables' relations for the initial conditions as well as the conditions emerged after the recommendations. All three fundamental diagrams (Figures V – VII) demonstrate the relationships between traffic flow properties, namely mean vehicle flow, mean density and mean speed, as they emerged from the simulation based on aggregated measurements of all edges for the 10 replications. Results indicate that the implementation of self-aware driving suggestions although it leads to safer and less aggressive driving behavior for each individual, it does not improve the performance of the road network. More specifically, self-improvement is evident from the lower mean density values which indicates that vehicles keep greater distances from the leading vehicles. Additionally, lower speeds are also observed after the adaptation of the recommended accelerations with the difference from the initial conditions being more significant in the case of saturated network flow (Figure V).

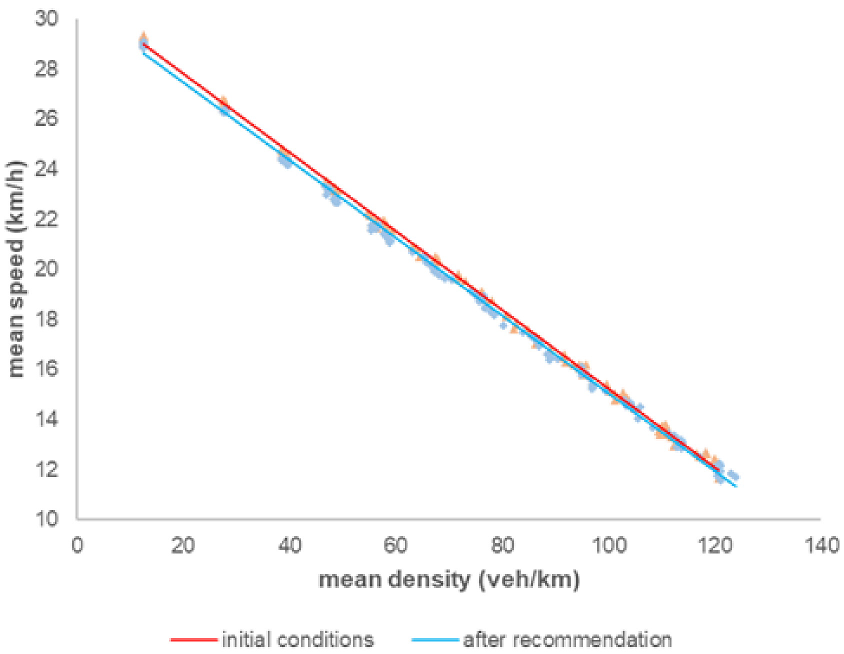


Figure V. Fundamental diagram of speed-density before and after driving recommendations, based on simulation results

Individual driving safety is augmented, yet the impact on traffic conditions is not similarly positive. The vehicles that move at lower speeds and with a lower density worsen traffic flow conditions, since fewer vehicles are served per time unit compared to the initial conditions. Nonetheless, this decrement of mean flow may be considered acceptable if assessed in conjunction with the positive effects on driving safety. However, based on the findings of this research, it can no way be concluded that the improvement of personal driving behavior is associated with a significant improvement in traffic conditions and therefore, the imposition of soft policy measures, such raising self-awareness with respect to individual driving safety and performance, it cannot be considered as a key measure for traffic management.

The fundamental diagram of flow-density seems to depict a uniformity between the initial and the final conditions, although some minor differences are observed with respect to the absolute value of capacity flow (Figure VI). Specifically, for the value of critical density, which was estimated 33.1 veh/km, the corresponding values of traffic flow are 360 veh/h and 358 veh/h for the initial conditions and after the recommendations respectively.

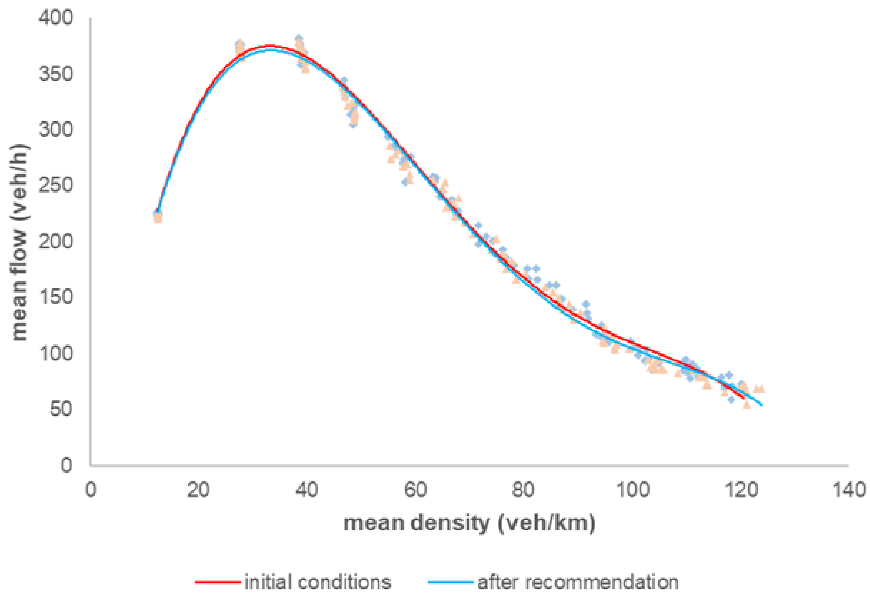


Figure VI. Fundamental diagram of flow-density before and after driving recommendations, based on simulation results

The flow-speed diagram is used to determine the speed at which the optimum flow occurs. For the initial conditions of the road network, the optimum flow occurs when vehicles move with 26.1 km/h, while the corresponding speed after the recommendation is reduced 3.4% with its absolute value estimated 25.2 km/h (Figure VII).

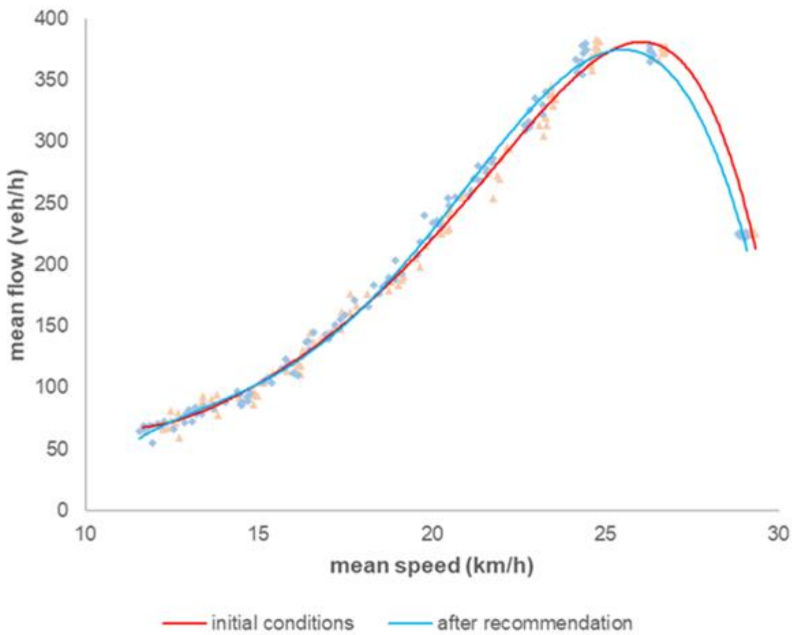


Figure VII. Fundamental diagram of flow-speed before and after driving recommendations, based on simulation results

Except for the performance of the network, another key performance indicator is safety. The assessment of the applied recommendations with respect to safety was performed by calculating the number of conflicts occurred between the vehicles during the simulation. Table

IX presents the number of conflicts that were observed for the entire traffic before and after the recommendation. There are three types of conflicts that can be identified from vehicles' trajectories, which are crossings, rear-ends and lane changes. Here, a special focus on rear-ends is given since the proposed recommendations only affect the car-following behavior of each driver.

Table IX. Safety performance indicators in Athens Network before and after applying driving recommendations

	Initial conditions	After recommendation [% difference]
Vehicles served	23,990	23,302
(in one hour of simulation)	*(27.88% of demand)*	*(27.08% of demand)*
Total number of conflicts	2.86 conflicts/vehicle	2.75 conflicts/vehicle [-4.2%]
Rear - ends	2.01 rear-ends/vehicle	1.90 rear-ends/vehicle [-5.5%]

A reduction of 4.2% of the total number of conflicts was observed when vehicles followed the corresponding driving recommendations, while the corresponding percentage of elimination for the rear-end conflicts is 5.5%. Although these percentages may not seem very high, the absolute number of conflicts that was calculated after the recommendation is significantly reduced by approximately 6,000 conflicts for the one hour of simulation. Rear-ends constitute about 33% of the total number of conflicts, which indicates that each driver gets involved in all different kind of conflicts during driving.

Some indicative results on the impact of the proposed recommendation system on emissions is provided. The corresponding Key Performance Indicator is the level of emissions for all different kind of air pollutants, namely Carbon Dioxide (CO_2), Carbon Monoxide (CO), Particulate Matter (PM_x) and Oxides of Nitrogen (NO_x). A significant reduction in all categories of emissions is observed compared to the initial conditions of the network, as shown in Table X. Findings revealed that the homogenization of acceleration profile for the entire traffic has led to a slightly reduced emissions per vehicle. Specifically, the reduction in all categories of emissions is estimated as follows: 2.5% in CO_2, 0.3% in CO, 1.3% in PM_x and 3.3% in NO_x. It should be noted that this improvement in the environmental conditions is very important since the proposed recommendation system had a positive impact on emissions despite the fact that the controller was not trained towards this direction.

Table X. Difference in vehicle emissions before and after applying driving recommendations

Emissions	Initial conditions	After recommendation [% difference]
CO_2	0.704 kg/vehicle	0.686 kg/vehicle [-2.5%]
CO	0.027 kg/vehicle	0.026 kg/vehicle [-0.3%]
PM_x	0.0133 g/vehicle	0.0131 g/vehicle [-1.3%]
NO_x	0.296 g/vehicle	0.287 g/vehicle [-3.3%]

Conclusions and main contributions

The main findings of the dissertation can be summarized in the following points:

- A two-level clustering approach can provide great insights on the characteristics that govern aggressiveness during driving and can be further exploited to distinguish safe from unsafe driving patterns.
- Six distinct driving profiles are able to describe the overall driving behavior that someone performs during their trip.
- There are two categories of drivers according to the average behavior of each driver resulting from how they drove in all their trips. In the first category drivers usually drive in a typical manner while in the second category drivers perform a number of unsafe driving actions or drive in an aggressive manner in the majority of their trips.
- The Actor-critic approach from the family of reinforcement learning algorithms can be exploited to find the best possible driving action for each dividual driver given the way they drove in their previous trip.
- When a controller provides driving recommendations to a fleet of vehicles, the acceleration profile of the entire fleet is harmonized on a value which is close enough to the acceleration decisions of a typical – safe driver.
- The application of a personalized recommendation system to a city's road network does not have a significant impact on traffic conditions.
- When each driver improves their own behavior, road safety is enhanced on the network. Specifically, critical conflicts between vehicles are significantly reduced after the application of the proposed system.
- The level of emissions for all different kinds of air pollutants is reduced which indicates that harmonization of the accelerations for the entire traffic can have an important positive impact on the environmental conditions.

Concluding, it should be noted that the deterioration of traffic may be considered acceptable if one takes into account the compensation through the benefits of adopting smoother driving behavior in road safety and emissions. To this end, policy makers and researchers should not neglect the real impact on all network's dimensions when planning traffic management strategies and applying soft and hard policy measures.

The present doctoral dissertation offers significant innovative contributions in five areas:

1. It makes use of an innovative naturalistic driving dataset. A great volume of data was available with high temporal resolution from real driving, enriched with a variety of factors that describe driving behavior, environment and other external attributes for each trip.
2. It proposes a methodological framework to extract driving profiles straight from the data, which describe the entire range of driving behavior. A data-driven approach is followed to classify critical driving patterns that appear during a trip by exploiting k-means clustering as being the most appropriate tool for this purpose.
3. It develops novel Reinforcement Learning algorithms to solve a real-world problem, this of assisting driving behavior. A deep Reinforcement Learning algorithm was

chosen as the most suitable tool to learn the optimal policy and suggest the appropriate action that leads to the best possible driving behavior for each individual driver.

4. It proposes a methodology which is capable of recognizing individual driving preferences and produce personalized driving actions to each driver. Specifically, an inclusive methodological framework is implemented which incorporates tools and methods that first recognize driving behavior of every user, then assigns every user to the corresponding RL controller version based on their overall behavior and finally produces personalized driving actions that mitigate aggressiveness and riskiness of driving.
5. It evaluates the large-scale network effects of implementing a personalized driving recommendation system on three areas of interest using specific KPIs, precisely on traffic, safety and emissions. Impact assessment of the proposed recommendation system is performed using a real-world scenario that of the Athens' Road network through microsimulation and by applying a before-after methodology to compare the values of the KPIs before and after the application of the system.

Limitations, impact and future research

As any other data-driven approach, this research as well, relied on some limitations with regards to problem setup and adaptation. Firstly, some limitations emerged from the need to match the RL output with the simulation properties. More specifically, one of the two components of the recommended action, the average acceleration of each driver, could not be imported into the microsimulation car-following model, which is parametrized by the acceleration ability of vehicles and therefore only the maximum acceleration is adopted within the simulation. Nevertheless, due to the nature of the phenomenon of driving, all parameters describing how a driver chooses to drive over a trip are inextricably linked with each other and therefore, the neglection of the average acceleration was not expected to have a significant impact over the results of the simulation. Besides car-following behavior, a driver during their trips takes actions regarding lane change, priority concession and other decisions concerning interactions with other road users. However, in this research the focus was explicitly on the car-following behavior as the ultimate goal was to create a user-centric system that looks only at the driver and does not require any external information from the road network in order to be trained and implemented. Thus, the proposed actions refer on the way the driver drives along the road, namely the way they choose to hit the acceleration pedal, which depends only on the personal preferences and perceptions of the driver. The lack of information about the environment can be considered as a limitation of the developed system, since its transformation into a context-aware system would give other perspectives both to the system itself and to the possibilities of its use as a traffic management tool.

An extension of the above limitation is the fact that since the system ignores the state of the environment it cannot operate real-time. In other words, the proposed methodology is not able to produce recommendations real-time, namely during a trip. Instead, an offline system is developed which suggests alterations on driving behavior in a sequence of trips for each driver. The integration of external information into the system would allow, at least conceptually, the real-time provision of driving recommendations.

Lastly, another limitation, which applies to all data driven approaches, is the generalization and transferability of the developed model and the corresponding outcomes. In most cases it is unclear whether the sample used to train the model is representative of the entire population and also whether its characteristics are similar to those of a different population. In this work, a big naturalistic driving dataset is used to develop the RL models which includes trips performed by a great number of drivers, nevertheless, it cannot be said that the results can be generalized and spatially transferred to another road network.

Besides the limitations described above, the outputs produced within this dissertation may have a significant impact on several aspects of both research (R), technology (T) and policy-making (P). Future research can benefit and significantly evolve by further examining the conclusions drawn with regards to the following points:

- (R) Aggressiveness does not necessarily constitute an unsafe driving habit and can be detected either as an individual behavior or in combination with other unsafe behaviors.
- (R) Reinforcement learning algorithms can be implemented in real-world problems and specifically, the DDPG algorithm can learn how to make human-like decisions on complex and high-dimensional environments.
- (R & T) The identified human driving profiles can provide great insights for human-like autonomous driving.

Technological advancements can be achieved in case the proposed recommendation system is incorporated in already developed software, such as insurance telematics apps and ADAS. Such system can be revolutionized, become more human friendly and adopt a more personalized way of supporting human decision making.

Moreover, policy makers could take advantage of the results of this dissertation to redesign soft policy measures and redefine the role of drivers in the current traffic management strategies, since in this work it was shown that the improvement of driving behavior on an individual level can have significant impact on road safety and emissions, but not a noteworthy impact on traffic conditions.

Finally, it can be understood that findings of this work can have far reaching implications for future research. Although this research provides significant contributions on driving behavior analysis, there is still much room in the exploration of driving behavior dynamics and thus, further research should be conducted in that direction involving enriched driving datasets and additional driving behaviors and parameters (e.g., cornering, tasks that cause distraction except from mobile usage). Moreover, the dedicated study of the dynamic evolution of driving behavior is also very important to provide answers to the question of how much and how rapidly driving profiles are altering over time. Another direction of future research concerns the recommendation system, which should investigate the way the produced recommendation should be passed to the driver in order for them to be understood by the user and then to be accepted by him. Furthermore, the identification of the required specifications that will enable the real-time operation of the system could also be a part of future research. Towards this direction, the most significant future research objective would be the modification of the proposed system in a way that it becomes context-aware, meaning that the system can interact with environment in which the agent takes decisions and have a full view of its dynamics and alterations. In this way, the proposed system could be implemented in real-time, and additionally it could also act as a traffic management tool which uses driving behavior as a key force of enhancing traffic efficiency.

1 INTRODUCTION

1.1 Background and motivation

The urban transportation system is changing because of the emerging innovations of Information and Communication Technologies (ICT), namely automation, connectivity and cooperation. In the near future, a variety of mobility solutions, spanning from traditional travel modes (private vehicles, public transport, etc.) to micro-mobility solutions (scooters, e-bikes, skateboards, etc.), sharing services (e.g., Mobility as a Service) and autonomous vehicles, should coexist on the road network while ensuring the effective mobility of road users and efficient use of road infrastructure. Furthermore, advances in information technology have facilitated the data exchange between road users and operators shaping the "informed-traveler" paradigm, an individual with multiple requirements regarding safety, comfort and level of service of the system (Nuzzolo *et al.*, 2014). In this complex environment, the role of traffic management becomes even more challenging since many transport operators with contradicting goals are involved and simultaneously there is a tendency towards a more human-centric approach of managing traffic.

The necessity to meet the personal mobility needs of each individual user within the context of a multimodal transport environment, has led to a distributed management system of decentralized transport. The management of such a system should treat each mobility service as a separate entity whose operation has to be ensured complementary and simultaneously with the seamless operation of the other units. Figure 1-1 graphically depicts the concept of decentralized traffic management in the era of connectivity and automation in transport.

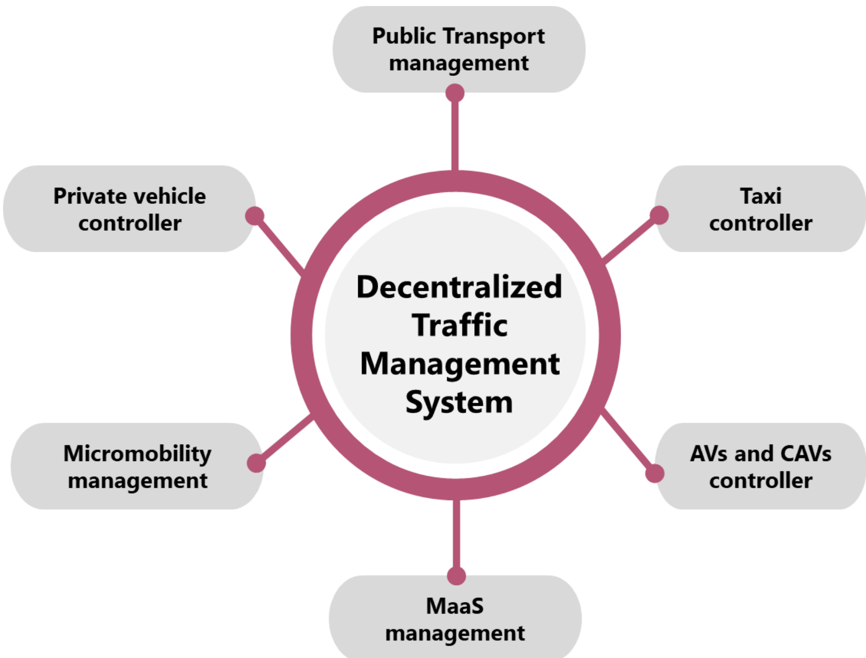


Figure 1-1. Schematic representation of the Decentralized Traffic Management Concept

Even in this ever-changing environment of connected, cooperative and automated transportation solutions, drivers are still the protagonists to fuel a safe driving environment (Vaiana *et al.*, 2014; Sagberg *et al.*, 2015) and a sustainable transport (Huang *et al.*, 2018). Beyond that, driving behavior affects traffic flow, fuel consumption, air pollution, public health as well as personal mental health and psychology. As a result, traffic management resolutions are mostly guided by understanding and improving driving behavior as well as mitigating vehicle use and ownership. Towards this direction, a variety of management measures have been applied over the years which can be separated in two main categories (Figure 1-2): (i) infrastructure interventions (hard measures) and (ii) modification of human behavior and actions (soft measures).

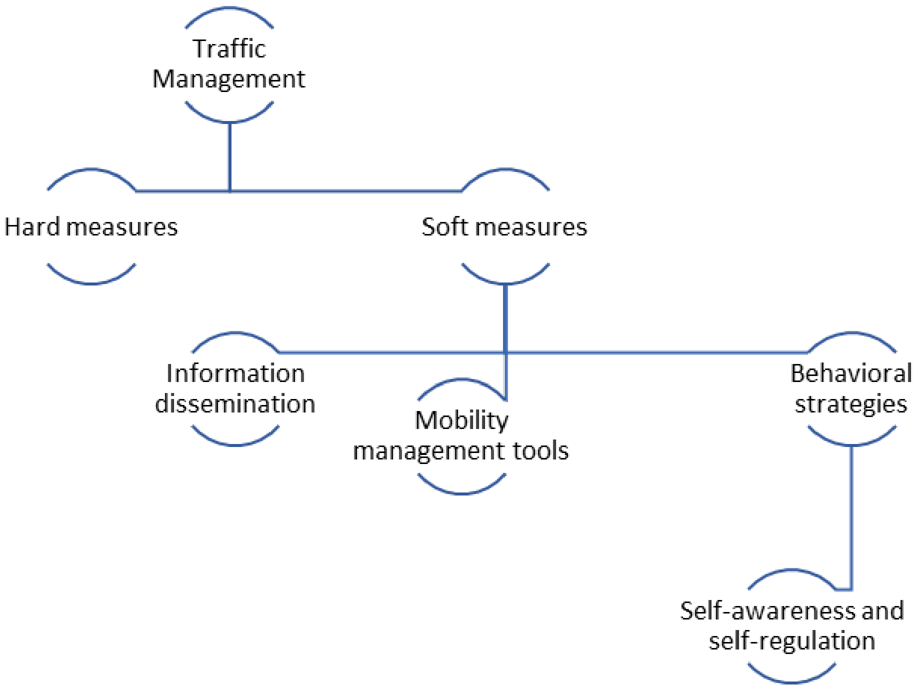


Figure 1-2. Categorization of traffic management measures

Hard transport management measures, such as congestion charging and other pricing policies that prevent car use, and efficient road space control, have been traditionally implemented as countermeasures of congestion and air pollution; yet, their results are not the ones expected due to public opposition and high financial costs (Gärling and Schuitema, 2007; Huang *et al.*, 2020). On the other hand, soft policy measures, also referred to as non-coercive measures, include psychological and behavioral strategies aiming at influencing the mindset of road users so that voluntarily change their travel behavior towards more sustainable mobility choices (Fujii and Taniguchi, 2006; Cairns *et al.*, 2008; Semenescu *et al.*, 2020). Soft policy measures are expected to reform the urban road landscape since their results have long-term effects not only on the instantaneous choices of drivers, but on the way they choose to drive in general (Möser and Bamberg, 2008; Bamberg *et al.*, 2011).

Nevertheless, humans in transport exhibit a typical selfish behavior, which means that in every situation they are expected to make those decisions that satisfy their needs, regardless of the consequences to the others or the system performance (Fehr and Fischbacher, 2003). In everyday travel, and especially while driving, man exhibits this selfish behavior, as their ultimate goal is to move as quickly and as comfortably as possible towards their destination. Many researchers have tried to change this selfish behavior and motivate the driver to think and behave as a member of the group, in the context of soft policy measures. However, the goal of each individual driver remains one "how to move as fast as possible". For this reason, a crucial question arises as to *how the needs of each driver will be met while at the same time the smooth and seamless operation of the road network will be ensured*. Within this context, the notions of self-awareness and self-regulations have been introduced to the design of behavioral strategies as a soft policy measure (Möser and Bamberg, 2008; Bamberg *et al.*, 2011), in a sense that the driver will become aware of the impact of their driving behavior, and thus, they will try to improve their driving performance on the road. However, until now, the concept of self-awareness has not been integrated in any of the already developed decision support systems for drivers.

Today's already applied soft policy measures include detailed travel information provision, road awareness campaigns, and marketing techniques focusing on personal travel behavior (Cairns *et al.*, 2008; Semenescu *et al.*, 2020). In addition, several mature driver assistance systems like the Adaptive Cruise Control, Lane Keeping Assistance System, Autonomous Emergency Braking have been successfully developed in order to ensure safety and comfort while driving (Meiring and Myburgh, 2015). Nevertheless, the main drawback of these systems is that they lack the personalized character; they are usually designed based on the behavior of an average driver and, therefore, may be found too conservative for aggressive drivers and too aggressive for the more passive drivers (Butakov and Ioannou, 2015; Tselentis *et al.*, 2016). Additionally, this lack of personalization restrains the potential of self-awareness in the sensitization of the individual regarding the impact that their decisions and behavior may have in others. As other human behaviors, driving behavior as well, is related with a variety of other issues such as congestion, road safety, interactions with other road users, air pollution and many more, which makes it a key component in the design and operation of the urban environment.

As a result, optimizing driving performance by addressing personalized aspects of driving behavior is a focal research area, which may have far reaching implications to traffic safety and operations, environment, as well as significant benefits for users (Vlachogiannis *et al.*, 2020). The first step towards this direction is the identification of driving behavior and the analysis of its dynamics. Driving behavior analysis is not a new concept. Over the years, a variety of methodological approaches have been applied in order to investigate the way drivers choose to drive (Chan *et al.*, 2019; Abou Elassad *et al.*, 2020; Mantouka *et al.*, 2020) by exploiting data from different sources, such as travel surveys, driving simulators, questionnaire surveys, GPS devices and only recently smartphone crowd-sensed data (Ziakopoulos *et al.*, 2020). Recent advances in cloud computing, Artificial Intelligence (AI) and Internet of Things (IoT) together with the high penetration rate of smartphones provide unprecedented capability to collect, exchange and analyze large volumes of heterogeneous data that enable the monitoring and understanding of mobility behavior, with a special focus on driving behavior for each individual (Tselentis, 2018). Consequently, this rapid technological process together with the storing and processing capabilities of today's smartphones, have paved the way for new research opportunities that include driving behavior monitoring, analysis and assistance. Several works have confirmed the efficiency and usefulness of crowd-sensed driving data collection schemes and their potential on driving behavior research (Araújo *et al.*, 2012; Kanarachos *et al.*, 2018).

Once driving behavior is being analyzed, and unsafe driving patterns are detected, the next step is the development of driving assistance systems that aim at improving drivers' performance, raising awareness on road safety and air pollution and improving driving experience. In addition to raising the awareness of drivers by providing feedback on the way they drive and the impact of their driving behavior, large scale studies with smartphones have shown that, when a driver is monitored, their behavior is relatively safer (Johnson and Trivedi, 2011). Nonetheless, as long as the main issue of the lack of personalization of the already developed driving assistance and recommendation systems still remains, it impairs the potential that behavioral strategies may have as a soft policy measure.

Within this context, the work contained in this dissertation is motivated by two main driving forces: i) the need to develop a driving recommendation system that treats each driver as an individual and proposes actions that meet his/her own driving preferences and, ii) the need to explore the actual impact of applying a personalized recommendation system on the road network. Following these research directions and taking advantage of the immense technological advancements, in this doctoral dissertation an inclusive driving recommendation framework is proposed which is able to recognize individual driving behavior and propose personalized optimal driving alterations towards safer driving.

1.2 Objectives

The main objective of this dissertation is to design a personalized driving recommendation system which is based on deep reinforcement learning algorithms and aims at enhancing driving safety through the mitigation of aggressiveness and other unsafe driving habits. Subsequently, the impact of controlling individual driving behavior is assessed with regards to network performance and road safety, as well as the levels of harmful emissions by properly adjusting parameters of traffic models in a city-wide scenario setting using microsimulation.

The above-described overarching goal of this dissertation can be divided in three major objectives as described below.

1. Exploit smartphone sensed data to understand driving behavior

Data gathered from smartphone sensors are exploited through data mining techniques, to identify some of the basic maneuvers and driving events. More specifically, machine learning approaches are used to detect aggressive behaviors such as harsh braking, accelerating and cornering events and identify driver's distraction by recognizing mobile usage while driving. In this dissertation, a large-scale naturalistic driving dataset which includes such driving events and other specific features are exploited in order to capture all different types of driving behavior (aggressive driving, distraction from the driving task, risk taking while driving, safe driving). Each specific type of driving behavior described by a set of variables is referred to as *driving profile.*

The driving profiles defined within this dissertation are universal, in terms that they cover the entire range of driving behavior from safe to aggressiveness and distraction during driving, and in addition they can be easily used to identify driving behavior of any single driver.

The first objective of this work is to develop a methodological framework for the identification of a specific number of driving profiles, that govern driving behavior on different levels (trip level, overall behavior) which can be easily transferable and interpretable.

2. Develop a traffic theory compatible personalized recommendation framework for improving driving behavior

A *recommendation system* is a system able to provide the most appropriate suggestions to the user for a specific task. Specifically, a recommendation system sequentially makes decisions on what to perform at the next step, based on the current available information (Tang *et al.*, 2019). Here, the aim is to develop a recommendation system for drivers which is able to suggest driving actions that improve individual driving behavior in terms of driving safety and aggressiveness. The actions proposed by the system are not decided arbitrarily, but result from the already observed driving behavior. In addition, recommendations are personalized; according to the United States National Education Technology Plan, the process of providing "*instructions in which the pace of learning and the instructional approach are optimized for the needs of each learner*" is referred to as personalized learning and has far reaching implications in recommender systems (United States Department of Education, 2017).

Therefore, the second objective is to design a recommendation framework using a deep reinforcement learning algorithm which is trained with naturalistic driving data so that to produce realistic driving actions that lead to improved driving behavior. The design of the RL algorithm enables the identification of the best driving policy for each individual which should be compatible with the traffic theory models that describe human driving.

3. Assess the impact of the recommendation system in traffic, safety and emissions

Several researchers have implied that the improvement of driving behavior would have significant positive impact on the performance of the road network. Nevertheless, no evidence has been provided towards this direction. To this end, the third objective of this dissertation is to investigate and quantify the actual impact of improving each driver's behavior, on three aspects: traffic, safety and emissions. For this purpose, the developed personalized recommendation system will be exploited in a simulation setting and several Key Performance Indicators (KPIs) will be estimated.

1.3 Innovation aspects

The concept of driving assistance systems is not new. Technological advances in this field are widely used in modern passenger cars, offering services such as adaptive cruise control, lane keeping assistance, driver drowsiness detection and many more. In this dissertation we are not concerned with developing systems that neglect the driver and apply predefined actions for safe driving. On the contrary, here, an innovative user-centric recommendation system is proposed which aims at improving driving behavior by proposing naturalistic driving actions that are already observed during real driving. To the best of author's knowledge, this is the first time that a comprehensive driving recommendation framework is developed, which employs a structured methodology from driving behavior identification, to recommendation provision and impact evaluation. This dissertation offers significant innovative contributions on five basic pillars (Figure 1-3):

1. **Data:** Makes use of an innovative naturalistic driving dataset
2. **Driving profiling:** Proposes and implements a data-driven driving profiles identification framework based on advanced machine learning
3. **Methods:** Implements state-of-the-art RL algorithms to solve a real-world problem, which is described be complex and continuous state and action spaces. Makes use of the DDPG algorithm for the first time in the context of driving recommendation provision system.
4. **Recommendation:** The system produces rational and personalized driving recommendations, that correspond to actual, already observed driving behavior, and, therefore, they can be easily adopted. Recommendation provision promotes self-awareness, contrary to ADAS which do not provide the opportunity to drivers to become aware of their erroneous driving habits.
5. **Impact assessment:** Evaluates the large-scale effects of implementing a personalized driving recommendation system on three areas of interest using specific KPIs.

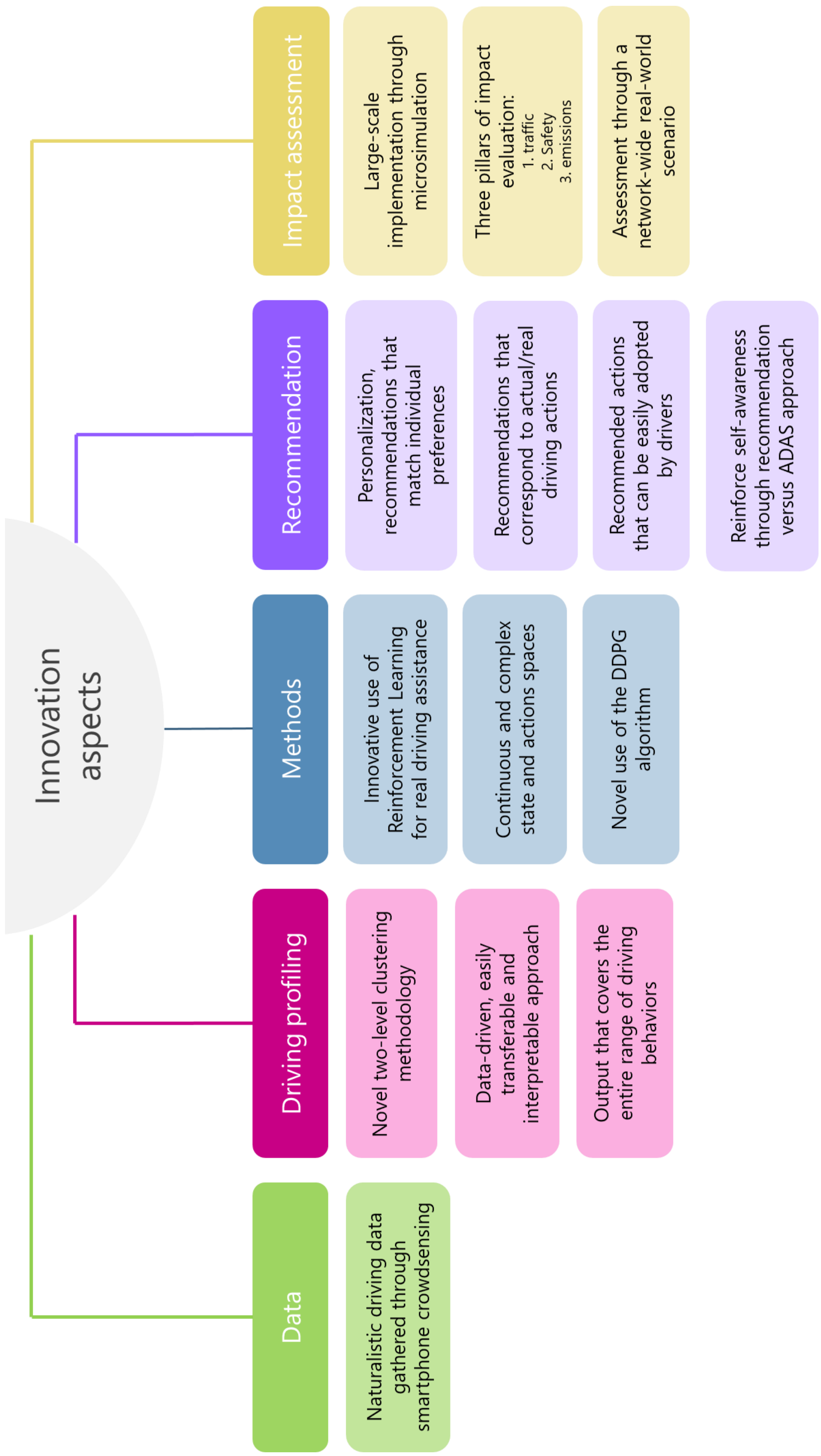


Figure 1-3. Innovation aspects of the doctoral dissertation

The data used here were collected using an innovative approach that is based on a smartphone application and thus, the dataset itself constitutes a novelty in the field of driving analytics. Contrary to previous research, real human driving patterns are identified by exploiting naturalistic driving data gathered from the crowd, and not through data emerging from driving simulators or controlled on-road experiments. Furthermore, the dataset includes driving behavior indices, such as harsh accelerations, decelerations and cornering events, metrics of mobile usage while driving and a plethora of statistical measurements of speed and acceleration in different road environments which contribute to the identification of a variety of driving behaviors.

On top of these, this doctoral dissertation contributes towards the understanding of the characteristics that govern the entire range of unsafe and abnormal driving behaviors, by exploiting real driving data collected from the crowd. Unlike the methodologies already applied in the literature, we do not use any predefined thresholds to separate the various driving behaviors, but instead, a data driven unsupervised learning approach is followed to define different driving profiles. In terms of the methods used, although clustering is a widespread methodology for unsupervised learning, its application in driving analysis is very limited. In addition, the two-level clustering approach followed to identify specific driving profiles can be easily interpreted and transferred, which adds an extra added value to the proposed methodology.

Recommendation provision is a wide field of research in a great number of thematic areas, from news recommendations, to online shopping, travelling and music. In the field of driving assistance systems only a few attempts have been made to develop a recommendation system for drivers, which are mostly based on universal driving actions applied to all users, as they emerged from rule-based and predefined threshold approaches with respect to safety and eco-driving. An innovation aspect of this dissertation is that it proposes and develops a personalized recommendation system, meaning that each individual driver's characteristics are identified in first place, and therefore, the suggested driving actions properly match the driving style of each driver. The dimension of personalization of the proposed system is very important for two main reasons; first, it increases the probabilities of adopting the recommendations by the drivers and secondly, it promotes the notion of self-awareness in the sense that the user becomes aware of their own unsafe driving habits and their impact, and thus, all adopted changes in their driving behavior are expected to have long-term positive effects. Additionally, a novelty of this dissertation is that it trains the system which produces the recommendation by exploiting naturalistic driving data. In this way it is expected to achieve higher rates of acceptance and adoption of proper (safer) driving behavior since the proposed actions correspond to actual driving habits.

Furthermore, another innovation aspect of this thesis is the application of state-of-the-art Reinforcement learning algorithms for assisting real driving behavior. Despite the fact that the unique ability of a reinforcement learning agent in learning from receiving a reward over different states of the environment without any training data makes it a perfect match for numerous recommendation problems (Afsar *et al.*, 2021), only a few applications have been made and even less in the transportation field. The development of a recommendation system for real driving by exploiting RL algorithms constitutes a novel methodological approach since driving is a complex, continuous task which can be described by a variety of parameters and the recommended driving actions belong to a large continuous space as well. In addition, the implementation of the DDPG algorithm is also considered as a novelty of this work, due to the fact that very limited applications of the specific algorithm exist in the literature and none of them in the field of conventional cars' driving assistance systems (Haydari and Yılmaz, 2020). The original DDPG algorithm (Lillicrap *et al.*, 2016) was adjusted accordingly, to learn the optimal driving policies for drivers with contrasting driving habits including the way they choose to accelerate and decelerate, the distance they keep from the leading vehicle and whether they drive above the speed limit (speeding).

One final niche innovation of the present research is the quantification of the impact that the application of a personalized driving recommendation system would have. To the best of the author's knowledge, this is the first time that the actual impact of such a system is quantified through a number of KPIs which can accurately describe the differences emerged after the recommendations in three basic pillars of urban networks: traffic, safety and the environment. On the top of this, the impact assessment methodology is implemented in a large-scale microsimulation scenario which corresponds to the Athens' Road network.

1.4 Structure of the dissertation

The remainder of the dissertation is organized in 6 chapters which are briefly described in the following.

Chapter 2 conducts an in-depth review of the literature in driving behavior analysis leaning on three basic pillars: the concept, data and the methods that were used in each study. Furthermore, it critically discusses the challenges that arise during data collection and storage, data preparation and data mining, modelling of driving behavior and decision-making and recommendation systems for drivers. Finally, it results in the identification of the existing knowledge gap in literature with respect to methodological and conceptual limitations of existing studies in driving behavior identification and driving assistance systems, and setting the key research questions for the present doctoral research.

Chapter 3 presents the methodology implemented to achieve the objectives of this dissertation and is divided into three main sections: (i) the thorough description of the methodological steps followed, (ii) the presentation of the theoretical aspects of all the machine learning methods used, as well as the basics of traffic flow theory and macroscopic fundamental diagrams, and (iii) the presentation of the naturalistic driving dataset exploited in this research.

Chapter 4 discusses in detail the implementation of the methods. Specifically, the methodology for recognizing driving behavior is applied, and the different driving profiles that can describe each driving behavior at trip-level are derived. As a next step, the average driving behavior of each driver is estimated and drivers are grouped into groups with similar characteristics. In the second section of this chapter, the conceptual design of the recommendation system is described and details regarding how the problem is structured based on the idea of Reinforcement Learning are provided. Finally, in the last section of this chapter the simulation scenario setting is thoroughly presented together with the key performance indicators that will be used to assess the impact of the proposed system.

Chapter 5 presents the results obtained by all the methods applied. First, driving recommendations emerged from the developed RL controllers are shown and critically discussed to provide insights regarding the differences between the outputs of these controllers. Then, findings obtained from the microsimulation are presented, with a special focus on the estimated impact of applying the driving recommendations in terms of traffic conditions, safety and harmful air pollutants.

Chapter 6 provides the conclusion of this thesis, an overview on the most critical findings and summarizes the major contributions of this dissertation. In this chapter, the limitations of this research are also discussed together with the impact of this work Finally, this chapter suggests proposals for advancing the present work further, along with other interesting lines for future research.

2 LITERATURE REVIEW

In this section, a review of state-of-the-art approaches that are proposed in the literature is provided, with respect to the three basic pillars of understanding driving behavior:

1. The data that are exploited, and data collection limitations, with a special focus on smartphone crowd-sensing.
2. The methods that are used, spanning from traditional statistical analysis to machine learning and more recently, reinforcement learning approaches.
3. The application of the results of driving behavior analysis to a variety of driving assistance systems.

The majority of works pertinent to the topic of this doctoral dissertation are reviewed, and their outcomes are critically discussed in order to outline current practices of driving behavior analysis and describe gap of knowledge that prescribe future research directions.

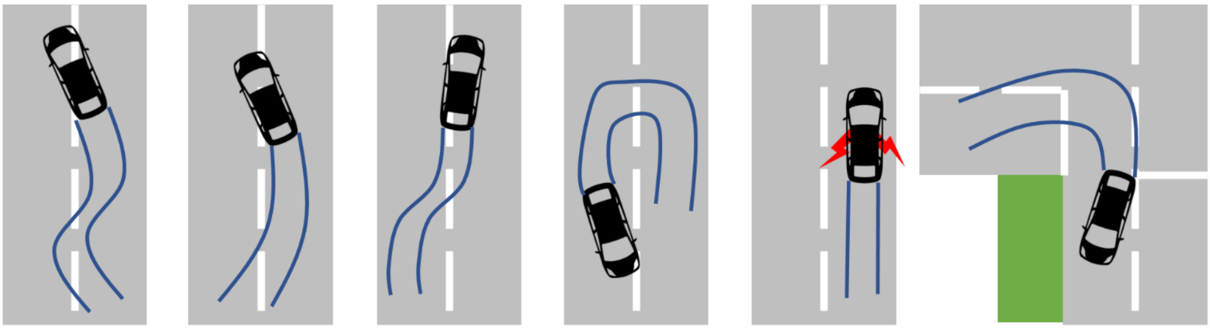
2.1 The importance of understanding driving behavior

Many definitions have been given for the notion of "driving behavior" or "driving style" which can be found in (Sagberg *et al.*, 2015). Here we embrace the definition given in (Lajunen and Özkan, 2011): "*Driving style concerns individual driving habits- that is, the way a driver chooses to drive*". Several driving profiles have been identified in the literature with regards to traffic and road safety since drivers differ in the way they choose to accelerate and decelerate, the distance they keep from the leading vehicle and whether they drive above the speed limit (speeding) (Miyajima *et al.*, 2007; Mantouka *et al.*, 2019). Among the most widespread driving profiles are:

- *Aggressive driving:* tailgating, harsh accelerating, braking and cornering, improper lane changing and many more (Tasca, 2000; J. H. Hong *et al.*, 2014; Smith *et al.*, 2016; Kockelman and Ma, 2018).
- *Distracted driving:* texting, eating, drinking or talking on the phone, where driver loses the focus on the driving task (Chen *et al.*, 2015).
- *Risk taking:* driving with excessive speed, violating traffic rules or driving too close from the leading car (Simons-Morton *et al.*, 2011).
- *Eco-driving:* driving in a fuel-efficient way, thus, minimizing pollutants (Andrieu and Pierre, 2012; Mensing *et al.*, 2014; Fafoutellis *et al.*, 2021)
- *Safe driving:* normal low risk driving behavior (Fazeen *et al.*, 2012).

The identification of these driving profiles relies on the detection of abnormal driving patterns, namely driving maneuvers that stray from the typical behavior and constitute a leading cause of serious traffic accidents. Based on the results of previous research, there are six types of abnormal driving behaviors as shown in Figure 2-1 (Yu *et al.*, 2017):

- Weaving: driving alternately toward one side of the lane and then the other, i.e., serpentine driving or driving in S- shape
- Swerving: making an abrupt redirection when driving along a generally straight course.
- Sideslipping: driving in a generally straight line, but deviating from the normal driving direction.
- Fast U-turn: a fast turning in U-shape, i.e., turning round (180 degrees) quickly and then driving along the opposite direction.
- Sudden braking: when the driver slams on the brake and the vehicle's speed falls down sharply in a very short period of time.
- Turning with a wide radius: turning cross an intersection at such a high speed that the car would drive along a curve with a big radius, and the vehicle sometimes appears to drift outside of the lane.



(a) Weaving (b) Swerving (c) Sideslipping (d) Fast U-turn (e) Sudden braking (f) Turning with a wide radius

Figure 2-1. Six types of abnormal driving patterns (Yu et al., 2017)

Another serious driving maneuver, which in most cases is neglected from relative research due to the inability of researchers to explicitly define it, is harsh acceleration. Harsh acceleration, as well as other harsh events such as braking and cornering, are significant indicators for driving risk assessment, and risk level correlation and classification (Bonsall *et al.*, 2005; Gunduz *et al.*, 2018). Fazeen *et al.* (2012) mentioned that the way a vehicle is maneuvered on the road can influence how other drivers react as they usually tend to follow previous movements to potentially avoid an unforeseen road hazard.

Due to the stochastic nature of driving, understanding and modeling driving behavior constitute a challenging topic for today's research. It is widely accepted that driving behavior vary between drivers according to a variety of characteristics that include sociodemographic (e.g., age, gender, ethnicity), driving experiences, emotions, and so on (Oltedal and Rundmo, 2006; Lin *et al.*, 2014). But besides that, even for the same driver, driving behavior may alter from trip to trip, or from situation to situation (Angkititrakul *et al.*, 2009; Lin *et al.*, 2014). Therefore, due to the complexity of driving behavior, the existence of the appropriate data that can capture all driving behavior dynamics under different conditions is vital. Towards this direction, the use of crowd-sourcing gained a lot of attention as the main source of data since driving information emerge directly from the crowd and can reveal the real dynamics of the phenomenon of driving.

The understanding of driving behavior is very important especially when it comes to the identification of the conditions under which a driver exhibits an unsafe driving behavior. Human factors, such as driving over the speed limit, distracted driving, driving under the influence of alcohol (Petridou and Moustaki, 2000; Sagberg *et al.*, 2015), are considered to be one of the main causes of road traffic accidents and therefore it is worthy to quantify the influence of driving behavior on crash risk (Tselentis, 2018). Despite mitigating safety, the detection and elimination of abnormal driving may also have implications on fuel savings, and consequently, result in lower accidents, less emissions and reduced operational costs for the driver (Van Mierlo *et al.*, 2004; Ferreira Júnior *et al.*, 2017). Once driving behavior can be detected and especially, the identification of unsafe and abnormal driving characteristics, drivers can receive feedback on the way they drive and even become aware of the impact on their driving habits. Therefore, advances in driving behavior analysis led to the flourishing of a large number of applications that aim at improving driving behavior and help drivers adopt more efficient and safer driving habits.

Finally, unsafe driving behavior detection systems are exploited in Insurance telematics market to sell usage-based insurance schemes (Tselentis *et al.*, 2017; Wahlström *et al.*, 2017; Geyer *et al.*, 2019). The evolution of conventional to usage-based insurance (UBI) can have far reaching implications on driving performance and road safety. Usage-based insurance pricing schemes refer to the formulation of pricing policies in accordance with the way the drivers drive, contrary to traditional insurance schemes where users are charged a lump sum according to the current pricing policy. This has been considered for long unfair and inefficient and therefore Usage-based insurance flourished alongside driving behavior analytics and driving monitoring advances. There are two types of UBI schemes (Liu *et al.*, 2017; Tselentis *et al.*, 2017):

- Pay-As-You-Drive Systems (PAYD): charging premiums are based on total exposure characteristics such as mileage and road network used.
- Pay-How-You-Drive (PHYD): charging premiums are based on individual driving behavior measuring parameters such as speed, aggressiveness, inattention etc.

The understanding of driving behavior requires the availability of naturalistic driving data. In addition, the application of UBI schemes requires the continuous monitoring of users as well as the collection of data describing how they drive. The main data source of naturalistic driving data is mobile crowdsensing, namely the collection of raw data from the sensors embedded in smartphones straight from the crowd. The next section provides a thorough review on the process of collecting driving data through crowdsensing.

2.2 Driving data collection through crowd-sourcing

2.2.1 The concept of crowd-sourcing

Collecting mobility data is essential to researchers and transportation planners in order to detect urban mobility and develop effective strategies to move towards more sustainable transportation modes like walking, biking and public transit (Jariyasunant *et al.*, 2012). Furthermore, the collection of naturalistic driving data is also very important to disentangle the complicated process of decision-making while driving, which incorporates a variety of parameters such as vehicle dynamics, personal preferences, road traffic and geometry, etc. Until today, driving data were collected using GPS devices and on-board diagnostics systems which offer the opportunity to gather naturalistic driving data instead of using outdated travel data collection methods like travel diaries, questionnaires etc. (Bricka *et al.*, 2009). A thorough review of recording methods and tools for driving behavior is given in (Ziakopoulos *et al.*, 2020). The rapid advances in mobile and communication technologies, more of the disadvantages of the previous data collection methods were eliminated. Nowadays, smartphones are carried by commuters all day long and in conjunction with the variety of sensors they are equipped with, they constitute a significant source of driving and mobility data. Smartphone sensors can be categorized into three groups according to their application in travel data collection as follows (Abdulazim *et al.*, 2013; Castignani *et al.*, 2015):

- Motion sensors:
    - Accelerometer, measures the device linear acceleration
    - Gyroscope, measures the angular rate of change (i.e., rotation velocity)
    - Magnetometer (i.e., compass), measures magnetic field strength
- Location sensors:
    - Global Positioning System (GPS) which is commonly used in outdoor settings
    - Network-based location services which use cellular network and Wi-Fi to determine the location (i.e., via triangulation)
- Ambient sensors:
    - Light sensor
    - Microphone
    - Proximity sensor, which detects nearby objects and can indicate when the phone is near the user’s ear (e.g., during a call)

Due to these components, smartphones have been adopted as useful tools to sense and compute data. Thus, the high penetration rate of smartphones has given another impetus to data collection, which is now done with high speed, frequency and accuracy. In this way, large-scale naturalistic driving data can be collected. Over the last decade, researchers take advantage of the great amount of data emerging from smartphone sensors, to investigate driving behavior, detect unsafe driving habits and predict road network conditions. The collection of data straight from the crowd is usually referred to as “crowd-sourcing”.

The word "crowd-sourcing" was coined by Jeff Howe in 2006, and was defined as *"…the act of taking a job traditionally performed by a designated agent and outsourcing it to an undefined, generally large group of people in the form of an open call"* (Howe, 2006). The power of crowd-sourcing is that it can bring massive intelligence to solve problems at an affordable price. Some tasks that are difficult for computers or individuals can be solved efficiently by crowd-sourcing to a massive group of people, including image tagging, audio translation, and so on (Yang *et al.*, 2015). A survey of existing transportation systems which use crowd-sourcing reveals that the predominant purposes of using crowd-sourcing in these projects are either data or feedback collection from the users (Misra *et al.*, 2014). Due to the wide penetration of smartphones at everyday life, mobile crowd-sourcing has gained the attention of many researchers. Figure 2-2 presents an overview of a crowd-sourcing system, where initially the required sensing task is introduced to the user, who provides the relevant data while his driving behavior is being monitored. Subsequently, data pre-processing techniques are applied and the user's driving behavior is analyzed based on specific mobility patterns. Then, results are sent back to the user containing feedback and incentives for an improved driving behavior.

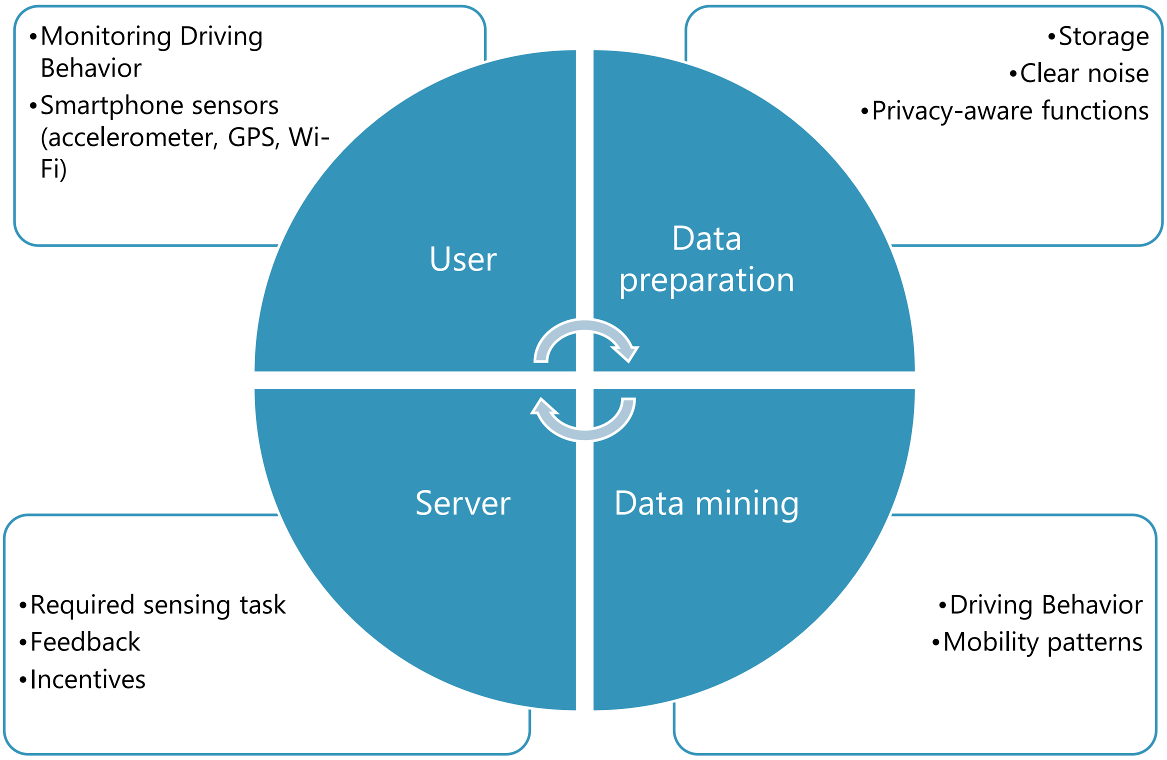


Figure 2-2. The main components of a crowd-sourcing system

The crowd of mobile users who accept and participate in crowd-sourced sensing tasks, is called the "sensing crowd" (Yang *et al.*, 2015). Since most modern smartphones are equipped with various sensors, many applications and platforms are developed to collect sensor data, a method known as "mobile crowd-sensing" (Chang *et al.*, 2016). A key characteristic of mobile crowd-sourcing is whether the crowd's contribution is participatory or opportunistic. In the case of participatory crowd-sourcing, computations are performed and data are generated by users, while in opportunistic crowd-sourcing data are generated from sensors and computations are performed automatically by the crowd's devices (Chatzimilioudis *et al.*, 2012).

Despite the wide range of advantages of mobile crowd-sourcing, on collecting driving and mobility data, there are some significant issues and concerns which should be adequately addressed in order for mobile crowd-sourcing systems to reach their full potentials. In the following sections, some of the key challenges are discussed when gathering such data from a smartphone.

2.2.2 Quality Issues

There are two main factors affecting the quality of crowd-sourced data. The first factor refers to the reliability of the user. Participants may misunderstand the required task, make mistakes or even deliberately cheat the system, which can cause errors or bad results (Xintong *et al.*, 2014). This issue refers to the trustworthiness of the data. The second factor refers to the technical characteristics of the mobile device that is used and may also affect the quality of the data.

2.2.2.1 Trustworthiness

The success of this method on collecting data relies on high level of participation from voluntary users. Unfortunately, the openness which allows anyone to contribute data, also exposes the process to erroneous and malicious inputs (Kanhere, 2013). The initial concern about quality of crowd-sourced data is the capability of the users to understand the requested task and supply the crowd-sourced system with all the essential and relevant information. Participants may not have the ability to perform the requested task efficiently and consistently, either because they cannot understand the concept of it or do not have the qualifications to perform the task (Wang *et al.*, 2017). Consequently, the captured data may be noisy by nature, and might require additional validation or scrutiny (Hsueh *et al.*, 2009). Moreover, users may deliberately be malicious and provide erroneous data to the crowd-sourced system (Kanhere, 2013). This kind of issue refers to the reliability of the sensed data. Finally, the challenge of data integrity is considered. The integrity of the data refers to the verification that the collected data is indeed from the users device and was collected at the claimed location (Mashhadi and Capra, 2011). In order to achieve integrity, sensing information must be representative of the user's behavior and habits.

2.2.2.2 Technical Characteristics

Despite users' intentions and abilities, the technical characteristics of the smartphones could also affect the quality of the data. The set of mobile devices, their sensing, computation, storage and communication capabilities may vary significantly (Louta *et al.*, 2016). Smartphones run on several different operating systems, most notably Android and iOS, which are frequently being upgraded, improving their features and capabilities, which should be considered during the sensing process, as they are directly related to the quality of the data.

Some devices have limited battery energy, low computational capacity and limited transmission bandwidth (Cao and Lin, 2017). Moreover, same type of data may be gathered from different sensors, e.g., location data can be gathered from GPS or Wi-Fi, and therefore have differing qualities. Another aspect that should be considered is that different types of data can be used for the same purpose, but with different quality and resource consumption trade-offs (Ganti *et al.*, 2011). Finally, even when same type of data is collected from the same sensor, different quality issues appear through the data due to the rapid change of the technical characteristics of the smartphones, concerning sensing accuracy, storage and computing resources and so on (Louta *et al.*, 2016).

An additional concern when dealing with crowd-sourced data is noisy data, which can be gathered concerning the way the user is using his/her smartphone. For example, ambient noise could be gathered if the smartphone is placed inside a pocket or in a bag (Xintong *et al.*, 2014). Furthermore, the same sensor may sense the same type of data under different conditions, if for instance the device is placed freely in a vehicle or is hand-held.

2.2.3 Battery Consumption

Even though most users charge their smartphones on a daily basis, significant increased battery consumption is an important issue when introducing crowd-sourcing applications. Energy is consumed in all aspects of crowd-sourcing platforms from sensing, processing and data transmission (Kanhere, 2013). The fact that battery life of smartphones is relatively short, limits the use of the devices for continuous sensing purposes (Birenboim and Shoval, 2015). Thus, it is hard to obtain accurate data in situations when continuous real-time data is needed. More specifically, different sensors deplete the smartphone battery in a different way while at the same time they are sensing the same type of data with differing accuracy. According to power measurements (Lin *et al.*, 2010) one can rank the sensors from the best to worse as follows:

- Battery consumption: (1) Wi-Fi (2) 3G (3) GPS
- Accuracy: (1) GPS (2) Wi-Fi (3) 3G

2.2.4 Privacy and Security

Westin (1967) gives the most relevant definition of information privacy: *"Information privacy relates to the person's right to determine when, how and to what extent information about him or her is communicated to others"*. Privacy in mobile crowd-sourcing is the guarantee that participants maintain control over the release of their sensitive information. This includes the protection of information that can be inferred from both the sensor readings themselves as well as from the interaction of the users with the participatory sensing system (Christin *et al.*, 2011). The sensed data may contain sensitive information of participants, such as identities, home or workplace location, mobile number, gender and so on. Without any suitable protection mechanism, smartphones are transformed into spies, capable to reveal such sensitive information (Kanhere, 2013).

There is a lot of research towards addressing the problem of privacy especially in location-based applications and in systems which detect mobility patterns. In those cases ensuring privacy is a challenge, either users explicitly provide information or information is implicitly captured (Nandan *et al.*, 2014).

Security issues can be divided into two categories: hardware risks and information security (Cilliers and Flowerday, 2015). The first category considers hardware risks, namely system viruses or malwares that may affect the hardware that is used to report information to the crowd-sourcing system. Furthermore, mobile devices can be stolen and are in general vulnerable to security breaches, both while sensing and transferring data. The second category has to do with the information that is reported to the crowd-sourcing system. It is stated that the participant has no control over the ownership of the information once it is reported to the crowd-sourcing system (Sarwar and Khan, 2013). This means that the collected data could be stolen, used for a different purpose than that originally agreed on, or made available to unauthorized parties. Another critical security issue is the lack of transparency concerning the physical location of storage, the security profiles of the system, ownership of the information and what can be done with it (Pearson, 2013).

2.3 Modeling Driver's Behavior

Understanding driver's behavior is key for improving road safety and optimizing the network's level of service. From statistical methods to advanced machine learning and computational intelligence, researchers have used a variety of methods to model driving behavior efficiently and accurately. Table 2-1 contains a detailed summary of crowd sensed data, smartphone sensors, methods and outputs of the most significant literature in the field of driving behavior analysis. Driving behavior analysis may be performed in two distinct levels; at the microscopic level, where the driver's behavior is analyzed both during a trip and how the same driver changes his behavior on successive trips, and at a macroscopic level, where the analysis aims to identify universal characteristics of each driving style using aggregated driving data taking into account a set of drivers. From statistical methods to advanced machine learning and computational intelligence, researchers have used a variety of methods to model driving behavior efficiently and accurately in all levels.

2.3.1 Statistical methods

Some researchers attempted to detect meaningful correlations among features that describe driving behavior. Paefgen *et al.* (2012) present a number of important statistical metrics and correlations between events recorded by a mobile and a reference IMU unit. In Chakravarty *et al.* (2013), the risk index of the individual driver is calculated at a specific time. The risk function in this case takes into account the number of risky and abnormal maneuvers such as sharp cornering, harsh brake and acceleration and hard bumps. Bejani and Ghatee (2018) applied statistical analysis on several driving features as an intermediate step between data preparation and more sophisticated driving behavior modeling. Another study has developed a mixed effects model to understand whether behavior indicators such as speeding, harsh maneuvering, harsh acceleration and harsh braking can stand as predictors of driver's distraction and more specifically if the driver uses his mobile phone when driving (Papadimitriou *et al.*, 2019). Findings revealed that exceeding speed and number of harsh driving events are negatively associated with mobile phone usage while driving, indicating that inattentive driving can be detected in the absence of other risky driving behavior.

Several researchers have applied a simplified threshold methodology in order to detect several abnormal driving events. In some cases, such methodology is also used to identify road anomalies such as bumps and potholes (Fazeen *et al.*, 2012; Bose *et al.*, 2018). This detection approach mostly relies on fixed threshold values applied on accelerometer data. Experimental results have shown that the accuracy of accelerometer data highly depends on environment's characteristics (mobile position in the car, vehicle's conditions, road type etc.) which constitutes such detection methods non-flexible (Ouyang *et al.*, 2018). Due to this fact, there is a difficulty to defining universal thresholds on the sensors' data for the identification of risky driving events. In order to overpass such limitations, researchers turned their attention on the more promising machine learning techniques which are usually easy transferable and more robust to changes in the environment.

Pattern recognition approaches, such as Dynamic Time Warping (DTW) to recognize driving behavior also appear frequently in literature. DTW allows to group similar mobility patterns even though the corresponding elements in the two series are not exactly aligned with each other. First, for two given time series a grid is constructed and then distance between all elements is calculated. DTW algorithm estimates the best path through the grid which minimizes the total distance. Johnson and Trivedi (2011) proposed a novel system that utilizes DTW to detect aggressive turns, accelerations, braking and lane change events. In Engelbrecht *et al.* (2016) authors used DTW to detect driving events and then utilized a heuristic classifier to categorize such events as safe or reckless. Results of this method were then compared with a maximum likelihood approach and findings revealed that the latter performed better in classifying a variety of driving maneuvers. More recent studies have used DTW to classify lateral maneuvers exploiting fusion of gyroscope and gravity sensor to acquire angular velocity (Singh *et al.*, 2017). Although DTW is widely used due to its fast and easy implementation on comparing timeseries data, it is not easily transferable due to its high dependency on the predefined threshold values.

2.3.2 Machine Learning and Computational Intelligence

Although a large body of literature uses statistical analysis methods to investigate several driving behaviors, over the last decades machine learning approaches have gained ground in this field. Recent studies have thoroughly examined the variety of machine learning techniques used to identify driving behavior (Chan *et al.*, 2019; Elassad *et al.*, 2020). Nevertheless, the main outcomes of each study are also discussed here with the aim to highlight the variety of machine learning methods used to identify driving behavior.

ML techniques have been used in Bhoraskar *et al.* (2012) to identify bump and braking events, and despite the optimistic results, authors state the importance of filtering and machine learning techniques (k-means clustering and Support Vector Machines - SVM) for better harsh events identification. In Hong *et al.* (2014) a Naïve Bayes classifier was implemented to identify aggressive driving styles with an accuracy of 90%. In addition, a pattern matching algorithm was used in another study (Saiprasert *et al.*, 2013) that outperformed a rule-based algorithm both for longitudinal and lateral events, while in Saiprasert and Pattara-Atikom (2013) the same technique was used to identify abnormal speeding events. Fazeen *et al.* (2012) exploited the three-axis accelerometer to analyze driving behavior and detect road anomalies (potholes, bumps, uneven or rough road). Their classification system resulted in high accuracy especially for rough or uneven road recognition. In Saiprasert *et al.* (2017) fuzzy logic and a rule-based algorithm are adopted to detect driving behavior such as harsh acceleration and braking or aggressive steering) by exploiting accelerometer and gyroscope data. In Koh and Kang (2015) researchers used Gaussian Mixture Model (GMM) with periodogram method to classify driving behavior on a gradient from smooth to aggressive behavior with a special focus on the elderly drivers. Due to the sensitivity of their dataset (which refers to elderly drivers' behavior) they have highlighted the limited performance of GMM in accurately classify driving style. What is described as the main issue in most of these studies is that the variety of hardware (sensors and smartphones), weather conditions during data collection, positions of the smartphone (even if it is not fixed) etc. make these techniques difficult to be transferable and proposed as a universal solution.

Some other researchers aimed at recognizing driver's state by implementing classification algorithms. Specifically, in Yi *et al.* (2019) several classification algorithms were compared regarding their performance on identifying three distinct driving states: normal, drowsy and aggressive. Results indicated that Random Forest had the greater overall accuracy when compared to the others classifiers (K-nearest Neighbors, Decision Tree, SVM). In Vlahogianni and Barmpounakis (2017a) the MODLEM algorithm achieved the maximum accuracy compared to other classification algorithms for detecting harsh events utilizing data from smartphones' accelerometers. A two-step k-means clustering algorithm was developed in Mantouka et al. (2019) to initially distinct aggressive from non-aggressive trips and then, trips were further clustered with respect to driver's distraction and risk taking. The main modeling challenge is to handle imbalanced datasets as abnormal driving behavior appears less frequently than normal driving behavior. Additionally, it has not yet been clarified how one can define the ground truth; therefore, the comparison between different approaches and datasets cannot be universal.

Neural Networks were used in Meseguer *et al.* (2017) to identify the degree of driver's aggressiveness using speed and acceleration measurements. Eftekhari and Ghatee (2019) evaluated the performance of two classification algorithms (Decision Tree and Naïve Bayes) compared to that of a neural network with 3 hidden layers. Their findings revealed that the neural network outperforms the other methods in detecting driving maneuvers. However, since the computational power that is required for the training and the validation of the models is significantly increased, several tasks need to be performed offline.

As has been seen in different transportation related research attempts, ML techniques have significant advantages over statistical methods (Karlaftis and Vlahogianni, 2011). With the availability of massive smartphone datasets, it was seen that while statistical methods could provide a first insight on the datasets, more advanced techniques were required in order to design accurate and efficient solution to the different challenges. Such an example can be found in Predic and Stojanovic (2015) who developed advanced ML classifiers to detect harsh driving patterns and reported improved results when compared to classical methods of activity analysis from accelerometer data based on statistical metrics of standard deviation, entropy, energy, mean value, etc. Furthermore, as shown in the literature, researchers usually utilize statistical methods when single events are being detected or abnormal driving behavior is separated from safe driving. On the contrary, ML methods are used when the whole range of driving behavior is investigated and a number of different driving profiles are detected. In Chan et al. (2019) different approaches are compared in terms of classification accuracy. Here special emphasis is given on the indicators used as inputs for driving behavior identification and in addition, further information regarding the data collection process are provided. It should be noted that the comparison of the different approaches in terms of absolute measures of accuracy is avoided since there are still several challenges that can lead to completely misleading results.

Table 2-1. Summary of data and methods used for driving behavior identification

Authors	Year	Driving Indicators	Experiment	Data collection			Software		Sensors	Analysis			Method	Driving Behavior
				Smartphone	OBD	Other device	Android	iOS		Statistical Analysis	Supervised Learning	Unsupervised Learning		
Dai et al.	2010	Acceleration, speed, lane changing	√	√			■		Accelerometer gyroscope	•			Pattern matching algorithm	Drunk driving
Johnson & Trivedi	2011	Maneuver recognition rate (hard left and right turns, swerves, and sudden braking and acceleration patterns)	√	√				■	Accelerometer gyroscope magnetometer	•			DTW	"typical", aggressive
Araújo et al.	2012	Driving condition, Evaluation of the Fuel Consumption	√	√			■		GPS		•		Linear Classification	scale from 1 (very poor consumption) to 4 (very good consumption)
Eren et al.	2012	Maneuver detection (unsafe left or right turns, lane departures, and sudden braking or speed-up)	√	√				■	Accelerometer gyroscope magnetometer	•	•		DTW, Bayes classification	Risky, safe
Fazeen et al.	2012	Maneuver detection (acceleration and lane changes) and road conditions	√	√			■		Accelerometer GPS		•		Pattern recognition, Classification	Risky, safe
Paefgen et al.	2012	Accelerations, Braking, Turns	√	√		√		■	Accelerometer gyroscope and GPS sensors	•			ANOVA, Kruskal-Wallis tests	Event detection
Predic & Stojanovic	2012	Acceleration, speed		√					Accelerometer GPS				DFT algorithm	Unsafe events detection
Castignani et al.	2013	Speeding, steering, acceleration	√	√			■		Accelerometer GPS				Fuzzy Logic scoring	Calm/Moderate/Aggressive events
Chakravarty et al.	2013	Acceleration, cornering	√	√					Accelerometer GPS	•				Risk score
Saiprasert et al.	2013	Acceleration, speed	√	√			■		Accelerometer GPS magnetometer	•	•		Rule-based algorithm, DTW	Normal/Aggressive driving events
Saiprasert & Pattara-Atikom	2013	Speed, position, heading	√	√			■	■	GPS	•	•		Rule-based algorithm, DTW	Speeding events
Meseguer et al.	2013	Speed, acceleration, rpm	-	√	√		■		Accelerometer		•		Artificial Neural Networks	Aggressive, normal, quiet
Zhu et al.	2013	Speed, acceleration, deceleration	√	√					Accelerometer GPS	•		•	Statistical analysis, Clustering	5 at-risk levels
Bergasa et al.	2014	lane weaving/drifting, sudden longitudinal movements	-	√				■	Rear-camera microphone inertial sensors, GPS	•			Kalman filtering	Inattentive driving
Engelbrecht et al.	2014	Speed, acceleration, rotation rate	√	√					GPS, Accelerometer gyroscope	•	•		DTW, heuristic method	Normal, Aggressive

Hong et al.	2014	Acceleration, speed, deceleration, engine RPM, throttle position, and steering wheel movement	√	√			■		GPS, compass, gyroscope, accelerometer		•		Naïve Bayes classifier	Aggressive violator, calm non-violator, aggressive non-violator, calm violator
Castignani et al.	2015	Acceleration, braking, speed and steering events, weather conditions, time of day	√	√			■		Accelerometer gravity sensor, magnetometer		•		Fuzzy Logic, Kalman filtering	Event detection
Koh & Kang	2015	Acceleration	√	√			■		Gyroscope		•		Gaussian Mixture Model	Aggressive, Non-aggressive
Engelbrecht et al.	2016	Speed, Acceleration	√	√		√	■		Accelerometer gyroscope	•	•		Maximum Likelihood (ML) classifier, DTW	Normal/Aggres sive maneuvers
Ma et al.	2017	Speed, acceleration	-	√			■		Accelerometer gyroscope GPS microphone				Kalman filtering (data correction), Fast Fourier Transform and cross-correlation algorithm	speeding, irregular driving direction change, abnormal speed control
Meseguer et al.	2017	speed, acceleration, rpm, throttle position	-	√	√		■		Accelerometer		•		Neural Networks	Degree of aggressiveness
Saiprasert et al.	2017	Heading, acceleration	√				■		Accelerometer gyroscope		•		Pattern matching algorithm, rule-based algorithm	Aggressive driving events
Singh et al.	2017	Acceleration, speed	√	√			■		Accelerometer gyroscope, GPS		•		DTW	Harsh driving events
Bose et al.	2018	Acceleration, braking	√	√			■		Accelerometer GPS		•		Classifiers	Harsh driving events
Kanarachos et al.	2018	Acceleration, braking pedal position	√	√	√	√			Accelerometer		•		Deep Neural Networks	Braking behavior
Eftekhari & Ghatee	2019	Acceleration, lane changing, turning maneuvers	√	√			■		Gyroscope, accelerometer, magnetometer	•	•		Neural Networks, Naïve Bayes, Decision Tree, Fuzzy system	Safe and Aggressive behavior
Mantouka et al.	2019	Harsh acceleration, harsh braking, speeding, mobile phone usage	-	√			■	■	Accelerometer gyroscope, GPS			•	K-means clustering	Aggressive, risky, inattentive, safe, non-aggressive
Papadimitriou et al.	2019	Mobile phone usage	-	√			■	■	Accelerometer gyroscope, GPS	•	•		Mixed binary logistic	Inattentive driving (mobile phone usage)
Yi et al.	2019	Speed, acceleration	-	√					GPS, accelerometer, gyroscope	•	•		Decision trees, KNN, SVM, ensemble learning	Normal, Aggressive, Drowsy

2.4 Systems for assessing and assisting drivers

The extended knowledge in the area of driving analytics has allowed researchers and other practitioners to develop advanced applications for the assessment of driving behavior as well as driving assistance systems that aim at improving drivers' performance, raising awareness on road safety and air pollution and improving driving experience. A recent study that has gathered the most relevant driving assistance systems can be found in (Meiring and Myburgh, 2015). Here, we focus on the exploitation of mobile crowd sensed data for the development of such systems and the usage of smartphones as the only communication platform between the system and the driver.

2.4.1 Drivers' assistance and recommendation systems

Advances in driving behavior analysis resulted in the development of Advanced Driving Assistance Systems (ADAS). ADAS are gaining widespread interest since they constitute an innovative and user-friendly technology which is able to meet safety and ecology standards by providing real-time driving tips (Kaur and Sobti, 2017). Some of the most widespread ADAS are adaptive cruise control, navigation assistance and rerouting systems, lane keeping assistance and driver drowsiness detection. Such systems have been developed mainly to enhance road safety and secondary, to improve travel comfort and driving experience. Advancements in driving analytics coupled with improved wireless capabilities of today's devices have allowed the development of real-time ADAS. Specific research has focused on real-time driver distraction detection which is mostly identified on the basis of driving performance measurements, such as lane position and steering control (Liu *et al.*, 2016) as well as speed control measurements (Tornros and Bolling, 2005).

While most of ADAS focusing on ensuring road safety and reduce car accidents, in case where driving experience improvement or promotion of eco driving is out of question, researchers have developed driver recommendation systems. Contrary to the widespread adoption of ADAS, research on driving recommendation systems is still at a primary stage. In Magana and Organero (2011), researchers have developed an eco-driving recommendation system which first detects driving style from the point of efficient driving through OBD and smartphone sensors, then, uses Random Forest to classify driving behavior and provide useful eco-driving tips. Araújo *et al.* (2012) have developed an eco-driving coach to promote fuel consumption efficient driving. First, they identify driving behavior and vehicle's conditions and then, they recommend one of the most popular eco-driving tips such as "switch off engine", "shift gear earlier", "your acceleration is too high", "you are too aggressive on throttle".

They have also developed some fuzzy rules to determine driver's intention to follow each of the provided tips. Another study has developed a context-aware driving assistant system aiming at promoting fuel-efficient driving (Gilman *et al.*, 2014). Researchers identify aggressive driving behaviors and then layout them on a map together with traffic and weather conditions to investigate specific driving patterns. In this way they are able to make recommendations on how it would be more efficient to have driven to the specific route and then give advice about how to improve their driving performance in the future.

Such systems have been shown to improve driving behavior and lead in adopting smoother and safer driving habits. As discussed in Staubach *et al.* (2014) drivers receiving eco-driving recommendations have adopted driving behavior characterized by less harsh maneuvers and maintain constant speeds. Another study has highlighted the importance of introducing gamification aspects in the recommendation systems in order to gain the engagement of the driver with the system and improve even more their behavior (Magaña and Organero, 2014). The concept of gamification refers to the idea of including entertainment and game-oriented design approaches in originally non-game contexts, such as mobile crowd-sensing applications (Wells *et al.*, 2014). An effective gamification framework is based on a series of metrics to quantify the success on a predetermined goal (improve driving behavior, adopt eco-driving habits, reduce fuel consumption, etc.), which is achieved through a set of tasks.

2.4.2 Drivers' scoring systems

Over the last decades, in addition to monitoring and providing recommendations for the improvement of driving behavior, there is a trend towards developing scoring and behavior assessment systems. Researchers and app developers seek to create environments for healthy competition and comparison between people in order to increase their awareness regarding major issues with a view that the latter are going to improve their behavior. In this context, scoring and behavioral assessment methodologies as well as the creation of ranking and benchmarking techniques have been strengthened. As mentioned before, scoring points and gaining budges by completing specific tasks are the most common gamified aspects, in conjunction with leaderboards, achievements, gifts etc. In a gamification framework, users may also communicate with others, compare their performance and compete (Vlahogianni and Barmpounakis, 2017b).

Smith *et al.* (2016) have developed a systematic framework for scoring driving behavior, which consists of three dimensions: risk score, operational score and economy score basically for fleet management applications. DriveSafe is an app that detects when the driver is distracted from the driving task and assigns a score based on driving behavior (Bergasa *et al.*, 2014). Acceleration, braking and turning events as well as lane drifting and weaving are first recognized and then a score is assigned to each driver taking into account the number of the events performed and their intensity. Castignani *et al.* (2015) have used data fusion to detect risky and aggressive driving events based on a fuzzy inference system. Subsequently, the number of identified driving events per trip is coupled with weather conditions and time of day of the trip in order to provide a trip score which ranges from 0 to 100. In Araújo *et al.* (2012), they have developed a driver evaluation system for raising awareness of drivers towards fuel consumption and eco-driving. For this purpose, smartphone sensors are used to collect speed, fuel consumption and acceleration measurements which are then used to classify driving conditions based on predefined thresholds. Then, a fuzzy logic approach was used to evaluate driving hints performance and choose the most appropriate to show to the drivers.

Although scoring systems are considered as an efficient way to raise the awareness of the drivers, some researchers highlighted the importance of gaining real savings as an incentive to improve individual driving behavior. To this end, insurance charging systems are gaining a lot more attention over the last decade (Husnjak *et al.*, 2015).

One of the first sectors that used the advanced technologies of modern smartphones to monitor and evaluate driving behavior of their costumers was insurance companies. The latter have established several insurance policies in order to charge drivers based on vehicle use and driving behavior characteristics, which include Pay-As-You-Drive (PAYD) and Pay-How-You-Drive (PHYD) systems (Wahlström *et al.*, 2017). In Tselentis *et al.* (2017a) the most popular Usage-based insurance (UBI) schemes have been reported. As they highlighted, there is evidence that UBI implementation would provide motivation to drivers to improve their driving behavior and alternate their behavior by adopting safe and more efficient driving habits. Handel *et al.* (2014) have also extensively discussed opportunities emerging from structuring insurance schemes based on driving features gathered through smartphone sensors. The latter have highlighted the importance of consciously design Usage based Insurance since there is always the risk that the provided feedback or recommendation is perceived the wrong way. In Chiu *et al.* (2014) a methodological framework for the detection of driving exposure factors is presented in order for insurance companies to re-structure their pricing strategies based on the estimated crash risk.

2.5 Main challenges

Driving behavior analysis is a multidimensional problem which requires a step-based approach, from the collection of the essential data to the development of the appropriate models, to efficiently describe the dynamics that govern human behavior while driving. In this section, the most critical challenges involved in the process of understanding driving behavior through smartphone crowd-sensed data are underlined, while a special emphasis is also given on the caveats that should be considered, which most times are being disregarded. These relate to data availability and quality, representativeness, context-based knowledge extraction, pattern recognition and modeling, recommendations for behavioral change, as well as impact assessment and real time operation.

Challenge 1: Enhancing data representativeness, availability and quality

In mobile crowd-sensing systems, the user himself is one of the fundamental components and therefore, user's engagement with the system is critical. Most studies aim at identifying universal driving behaviors and determine general rules controlling unsafe driving behavior. Since the identification of driving behavior relies on data-driven approaches, the variability of the data is of great importance. The proportion of the target – population which is actually engaged in the data collection process pointedly affects data representativeness and, inevitably, affects data quality. In addition, the long-term involvement of the users is also crucial when driving behavior is analyzed due to the complexity of human behavior in relation to time and how slowly or not it changes over time and various stimuli. In the literature, several techniques have been determined as successful interventions for ensuring long-term user engagement with the crowd-sensing system. Some researchers have highlighted the importance of providing incentives to users in order to get engaged with crowd-sensing systems in the long run Yang *et al.* (2016). In Musicant and Lotan (2016) researchers have highlighted the effectiveness of group incentives in motivating drivers especially the younger ones. Although several types of incentives have been recently deployed, the degree to which they can raise the awareness and drive behavioral changes is still heavily under researched.

Further, behavioral variability in an available dataset is usually disregarded with significant implications in model generalization power. Researchers who tend to find large datasets as a testbed for exercising and evolving their machine learning skills, will realize that their models will soon become obsolete (rapidly decreasing accuracy metrics), as the dataset grows, mainly due to the fact that the initial sample is not representative of the users' characteristics. But, even if we try to control the statistical characteristics of our sample, the reality tends again to surprise us: first, most of the times researchers cannot have access to sample characteristics (e.g., users, age, gender), due to privacy limitations. Second, it takes much time to collect and process a large dataset especially if one is to ensure specific users' and behavioral representativeness. The above can be efficiently tackled by introducing processes of constant training and processes to address models' resilience to changes.

Data availability and quality are essential ingredients for every data driven approach met in transportation literature. Availability is not guaranteed for several reasons. First, data may be private or restricted to access. Ensuring users' privacy and data protection is of vital significance. Techniques to secure the system from the infringement of unauthorized parties is a main requirement. Lately in the EU under the strict rules of General Data Protection Regulation (GDPR), sensing data from the crowd became even more challenging especially in terms of data privacy and security. So far, researchers seemed to disregard the importance and implications of developing strategies for ensuring data privacy which is expected to change in the near future. Further, data may not be available at the desired resolution. It is a fact that certain driving phenomena (e.g., distraction due to smartphone interaction, lane changing etc.) may require very detailed recordings (e.g., 100Hz), whereas others can be easily observed in coarser levels. Identifying the proper smartphone sensing data resolution, in relation to the application developed, is heavily overlooked in literature and significantly affects the detection capabilities and the understanding of the driving task and context. But, even if someone chooses to get the most of the sensors' capabilities installed in smartphones (usually 100Hz resolution), this will probably lead to an unrealistic and non-sustainable data collection scheme due to battery drain. While increased sampling rate is desirable for improving the predictive power of the models, this can have negative effect on the user experience, since the rapid consumption of device's resources may discourage the users in contributing to the system as was identified in a previous challenge.

Data quality may have far reaching implications to the understanding and modeling of smartphone data due to uncertainties introduced by the fact that smartphone devices: i. are of various technologies, ii. can be placed anywhere on the user or in the vehicle and iii. have sensors that may record in different frequencies and asynchronously. Data reorientation and synchronization strategies that are usually applied to correct data and place them in a form ready for modeling require significant effort, but also do not ensure against faults and noise. Quality assurance schemes in the smartphone crowd-sensing framework are necessary especially for driving analytics, due to the fact that phenomenon detection relies on the extreme values and usually include noise filtering, data reorientation and geotagging.

Challenge 2: Identify the context from the data

User behavior is significantly influenced by the mode of transport, the type of road network, traffic control, weather conditions etc. To detect critical patterns, as well as understand and model the behavioral characteristics, the context of the data should be extracted from the data themselves, by jointly considering other external information. Dealing with the problem of context extraction from data has a significant added value to model resilience, since the quality of the features which can be extracted from the raw data primarily reflects the overall accuracy of the system (Ignatov, 2018). But, how should we understand the context through user agnostic experiments and blind crowd-sensing systems? In user agnostic environments, the main tools for deriving context from multisensory data identified in literature are: i. Data fusion, ii. Feature engineering and iii. User/behavior profiling.

Based on the well-known definition given by Hall and Llinas (1997) "*data fusion techniques combine data from multiple sensors and related information from associated databases to achieve improved accuracy and more specific inferences than could be achieved by the use of a single sensor alone*." There are three nonexclusive categories of data fusion techniques: (i) data association, (ii) state estimation, and (iii) decision fusion (Castanedo, 2013). The data association is mainly referring to the process of correlating several multi-sensor measurements about the object of interest with each other or in other words, it is about letting the data fusion process know which particular measurements are supposed to be fused to provide the essential information (Schmitt and Zhu, 2016). In state estimation the goal is straightforward the aligned and correlated measurements from multiple sensor sources must be fused in a well-defined estimation framework to infer the desired information about the target in an optimal way (Yan *et al.*, 2016). Finally, in decision fusion, the fusion task is not applied to the combination of data acquired by different sensors; instead, only a single input data set is used, and the fusion step is applied to several preliminary classification results obtained from this input (Fauvel *et al.*, 2006; Schmitt and Zhu, 2016). Data fusion is important when it comes to ensure that all available information can be jointly considered; for example, weather data with smartphone sensing and social media information can be fused to extract mobility patterns shifts due to weather changes or under differing travel purposes.

Feature engineering is very important when dealing with machine learning since all ML algorithms use some input data to create outputs. This input data in most cases refers to features which are numeric representations of characteristics, properties and attributes of the raw data (Zheng and Casari, 2018). Feature engineering can enable the identification of those features that are critical to context awareness. Different datasets require different feature extraction approaches based on the nature of the dataset. Nevertheless, the most well-known techniques of feature engineering are the following:

1. Imputation: it is used to handle missing values in a way that it preserves the data size, contrary to the easy task of dropping the entire rows or columns of the dataset that have missing values. Imputation can be either numerical or categorical based on the type of data. In the case of numerical imputation missing values are replaced either by 0 or a statistical measurement (e.g., median), while in the categorical imputation the most frequent value may be used.
2. Handling Outliers: Outlier detection is a very challenging task that should be performed before feature engineering using the most data-appropriate technique from great variety of techniques that are available (Wang *et al.*, 2019). Once outliers are detected, researchers have to use their intuition, perform experiments and provide some thoughtful discussion before deciding whether to exclude outliers from the dataset or not. It is understood that in the outlier detection process, it is significant to consider the context and the purpose of detecting the outliers.
3. Binning: The main motivation of binning is to enhance model's robustness and prevent overfitting. Nevertheless, binning data results in the regularization of them, and thus valuable information may be lost and model efficiency impaired (Shi *et al.*, 2020).
4. Logarithm transformation: The so-called log transform helps to handle skewed data and after transformation, the distribution becomes more approximate to normal.
5. One-Hot Encoding: The method of spreading the values in a column to multiple flag columns and assigns 0 or 1 to them. In other words, it transforms categorical data to a numerical format and enables the grouping of categorical data without losing any information (Seger, 2018).
6. Grouping Operations: This method is applied to datasets that do not match the "tidy" format. Based on the definition given by (Wickham, 2014):*"Tidy datasets are easy to manipulate, model and visualize, and have a specific structure: each variable is a column, each observation is a row, and each type of observational unit is a table."*. The key point of group by operations is to decide the aggregation functions of the features. For numerical features, average and sum functions are usually convenient options, whereas for categorical features the process is more complicated.
7. Scaling: With only few exceptions, ML algorithms don't perform well when the input numerical attributes have very different scales. When applying scaling techniques, the

continuous features become identical in terms of the range. There are two main types of scaling (Géron, 2017):

- Normalization (or min-max normalization) scale all values in a fixed range between 0 and 1.
- Standardization (or z-score normalization) scales the values while taking into account standard deviation. Although standardization does not bound values to a specific range, which may be a problem for some algorithms, is much less affected by outliers when compared to normalization.

Some examples of feature extraction can be found in Mantouka et al. (2019) where accelerometer signals were processed to extract features such as harsh accelerations per km and harsh brakes per km which are then used for driving aggressiveness detection. Another feature which is usually detected at first place is travel mode (Nikolic and Bierlaire, 2017; Efthymiou *et al.*, 2019). Mode detection techniques should be applied first and then data should be cleaned based on the detected mode, as the noise is not consistent between different modes due to different dynamics. For example, a smartphone on a motorcycle is affected significantly by the riders' maneuvers, while in public transport there are usually micromovements due to the device's position or its usage, as highlighted in the second challenge.

User profiling or behavior profiling refers to the process of detecting patterns in the data that can be clustered into groups which share common characteristics. In addition, user profiling enables the identification of users' preferences, choices and requirements based on the domain of interest, and therefore, user profiling helps personalization (Kanoje *et al.*, 2014). When studying driving behavior, user profiling refers to the categorization of drivers into groups based on the way they behave during driving. In accordance to (Ferreira Júnior *et al.*, 2017), driver behavior profiling is the process of automatically collecting driving data and applying a computational model to generate a safety score for each driver.

Challenge 3: Detect abnormal patterns and risky drivers

A common research line followed by most studies refers to the threshold-based approaches for detection of deviations from normal driving. The oxymoron lies in the fact that many studies conclude that no universal thresholds can be applied, since the technical variability of smartphones' sensors affects the extracted signals and of course the threshold of what is considered as abnormal. Indeed, threshold-based approaches are the first step to follow for detecting abnormal driving patterns especially in systems where no prior knowledge exists (e.g., annotated samples that can be used for supervised learning). However, a device-agnostic context agnostic approach to setting the thresholds is required to ensure that the latter are not affected by device micromovements, sensors' characteristics and device orientation and positioning. Even the consideration of a non-threshold methodology would be beneficial for the establishment of a universal driving behavior profiling framework.

Another commonly disregarded aspect in extreme behavior detection is how to generalize from trip-based characteristics to user specific profiles, meaning how much time a driver should be monitored so that his behavior is understood. This becomes critical in a macroscopic analysis level, where the system should provide recommendations to the user on how to improve his driving style and become a safer and more efficient driver. The answer to the question of how much time someone should be monitored so that the convergence of his driving characteristics is observed, is not unilateral. It depends on the user's perception and attitude on the road, but also on the roadway, traffic and control conditions. A recent study attempted to identify the essential amount of data when a single driving characteristic is examined (e.g., aggressiveness) (Stavrakaki *et al.*, 2020). Nevertheless, a methodological framework for estimating the amount of driving data that should be collected for each driver in order to acquire a clear picture regarding his overall driving behavior is missing from literature.

Challenge 4: Modeling efficiency, transfer learning and explainability

An efficient model is the one that is not only accurate, but also does not produce systematic errors. In highly volatile problems treated with data-driven techniques, achieving accuracy, but also taking care of overtraining and the properties of the error are of outmost importance (Karlaftis and Vlahogianni, 2011). For the case of smartphone sensed datasets that are usually imbalanced (abnormal driving or even accidents may be very rare in datasets), achieving to build and operate efficient models is a tricky and tedious process. While researchers turn to resampling techniques or generating synthetic datasets, different challenges arise when it comes to which technique is suitable for each problem. For example, should the dataset be resampled so that both classes are equally represented, or should one class be overproportioned to the other and if yes, how much? Although these questions have been addressed partially by the literature, enriching the existing datasets with more features could also provide significant improvements to deal with this challenge.

Given this lack of representativeness of extreme behaviors and accident conditions, the use of transfer learning seems also an appropriate pathway. Transfer learning and domain adaptation refer to the situation where what has been learned in one setting (e.g., the distributional characteristics of traffic volume in a single arterial) is exploited to improve generalization in another setting (e.g., arterial traffic volume in a different location) (Mairaj *et al.*, 2019). In simple words, transfer learning leverages the knowledge and pattern recognition capabilities developed based on another problem to facilitate the learning process of the problem at hand. The main limitation of ML models, and specifically deep learning models, is the time-consuming training process. Even when hardware with great computational capabilities is used, the training phase in the presence of big data cannot be neglected as a future challenge especially for real-time applications. Transfer learning may prove as a good solution for the time-consuming process of training an ML model for detection based on smartphone data.

ML models can easily tackle several data limitations such as noisy data, imbalanced datasets and the majority of them are transferable and resilient.

Nevertheless, building accurate models should not be the sole concern. Most researchers frequently follow the path of least resistance by comparing approaches based solely on model accuracy; however, there is a “thin line” between modeling accuracy, model simplicity, suitability and usability. Researchers should keep in mind that the produced models should be actionable, meaning simple to operate and maintain, accurate enough to produce reliable results and easily integrated to complex systems. Moreover, in cases where emphasis is given to the explanatory power of the models, modeling should clearly address the issues of causality. It is well known that correlations do not imply causation. Addressing causalities through ML modeling – either being for function approximation, pattern recognition or time series analysis - is not a straightforward process and several statistical constructs can be used to tackle this issue (Hlaváčková-Schindler *et al.*, 2007; Karlaftis and Vlahogianni, 2011; Lavrenz *et al.*, 2018).

Challenge 5: Raising awareness and changing attitudes

Although road safety is an issue of concern to researchers, practitioners and even drivers themselves, not everyone realizes how driving behavior is correlated with the levels of safety on the road. What is important is to make drivers aware of their driving behavior’s impact. Risk taking, aggressiveness and distraction from the driving task constitute the main reasons why a driver gets involved in a car accident (J.-H. Hong *et al.*, 2014; Dingus *et al.*, 2016). ADAS aim at supporting drivers while driving and preventing such unsafe behaviors. Another growing trend is coupling safety and sustainability of transport. Therefore, as already said, the aim of today’s driving recommendation and assistance systems becomes twofold: promote eco-driving while at the same time ensuring road safety. On the eco-driving side, changing driving attitudes seems to be easier since this driving style is correlated with monetary cost and fuel savings. On the other hand, regarding safe-driving, changing people’s unsafe driving habits seems more challenging especially because people tend to disregard the impact of their driving behavior. In this case as well, powerful incentives should be given since the ultimate goal is not only to motivate drivers to participate in a cause, but raising their awareness as well.

An additional concern is that driving recommendations systems premise continuous cooperation between the driver and the system. In such systems, the driver is constantly receiving driving tips and suggestions for improving their driving efficiency or even their driving experience. If the driver does not intend to accept the recommendation provided from the system, then consequently, the system loses its potentials of improving road safety, traffic conditions and driving experience. To this end, it is crucial to raise the awareness of the drivers regarding the impact of their driving habits as well as ensuring their long-term involvement with the system. Most researchers focus on the importance of education and training as key for changing drivers' attitudes (Zhao *et al.*, 2019). Nevertheless, recent trends require the continuous monitoring of driving behavior and the provision of recommendations and tips to be constant if not real-time.

Challenge 6: Lack of personalization

Personalization refers to the process of making something suitable for the needs and preferences of a particular individual (Hasenjager *et al.*, 2020). Within a framework of a system, personalization can be achieved either explicitly or implicitly. In the first case, the users are left to specify and customize the service they receive by themselves (explicitly), while in the second case, the system should be able to automatically detect and infer each users' specific needs and preferences (implicitly) (Hasenjager and Wersing, 2018; Ponomarev and Chernysheva, 2019). In driver assistance systems, personalization is a very key feature that can largely determine the acceptance of the system by the driver himself. According to previous studies, if the advanced driver assistance system (ADAS) cannot fit drivers' preferred driving behavior, the conflict between the driver and the vehicle will occur (Parasuraman and Riley, 1997). For this reason, the rationale for personalization in ADAS is to improve the driving experience by adapting the assistance system to the preferences and needs of the assisted drivers. Nevertheless, the concept of personalization has not yet been incorporated in already developed ADAS since the focus is mostly concentrated on safety and usability. The personalization of ADAS is not an easy task as great amount of driving data must be available, in order to capture different drivers' preferences, driving styles, skills and driving patterns, and other driving behavior analysis caveats should be first overpassed to identify different types of driving behavior to properly adapt ADAS vehicle control with driver's expectations (Hasenjager and Wersing, 2018). To this end, it is important that researchers and practitioners recognize the significance of the personalization aspect and put great efforts towards the development of personalized driving assistance systems.

Challenge 7: Real-time operation

Until recently, a very small body of the literature has developed driving behavior detection and recommendation algorithms that are able to operate in real-time. However, as the need for ensuring the resilience of nowadays transportation system through the management of traffic, road accidents and emissions has emerged, there is demand for ensuring real-time operation of driving recommendation systems. As mentioned before, driving behavior is a major contributor not only of road safety but road network conditions as well, and, thus, the chance to improve driving behavior can be feasible only by developing easy interpretable and responsive recommendation systems. Considering the complexity of human behavior as well as the previously described methodological challenges, future research should place a lot of effort in developing recommendation systems able to operate in real-time, especially if they provide personalized suggestions and tips. Advanced computational intelligence techniques are expected to have promising results and efficiently correspond to the challenging task of user-tailored real time recommendation systems.

Challenge 8: Moving from user-centric systems to network level management

Even in the case where all the aforementioned challenges were being addressed, there would still be a major concern: how to move from the development of user-centric driving assistance systems to a system-centric approach in the sense that, the impact on the network is taken into account alongside with the goal of improving driving performance and safety?

Traffic conditions can be studied in a microscopic level, where each vehicle is an entity that moves with each own characteristics and patterns, or in a macroscopic level, where aggregated measures of the characteristics of the entire traffic can be estimated. In traffic flow theory there are three macroscopic flow variables that reflect the average state of the traffic, namely flow, density, and speed. Relations between these traffic flow properties are usually graphically represented using fundamental diagrams of traffic flow, which can be used to predict the capability of a road network, or its behavior when applying inflow regulations and measures or speed limits. Macroscopic Fundamental Diagram (MFD) as a concept was first proposed in [15]. The theoretical form of these fundamental diagrams is illustrated in Figure 2-3. The MFD functions can aid agencies in improving network accessibility and help to reduce congestion by monitoring the number of vehicles in the network.

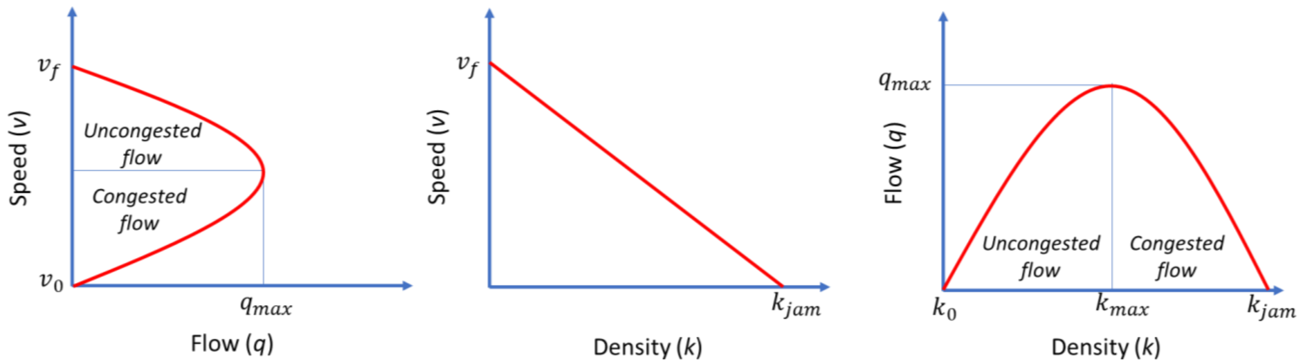


Figure 2-3. Macroscopic fundamental diagrams of traffic flow according to Greenshield

The speed – flow diagram is used to determine the speed at which the optimum flow occurs, and consists of two branches, the free flow and congested branches. In the first speed-flow diagram branch which corresponds to the uncongested flow, as the flow increases the speed decreases until the optimum flow q_{max} is reached.

The speed-density relationship is linear with a negative slope; therefore, as the density increases the speed of the roadway decreases. The diagram crosses the vertical axis at the free flow speed v_f, and the horizontal axis at the congestion (critical) density k_{jam}. The speed approaches free flow speed as the density approaches zero, while the speed reaches approximately zero when the density equals the critical density.

The flow-density diagram is used to determine the traffic state of a road network and has a parabolic form. The intersection of free flow and congested branches of the diagram is the apex of the curve and is considered the capacity of the road network, which corresponds to the traffic condition at which the maximum number of vehicles can pass by a point in a given time period. The flow and density at which this point occurs is the optimum flow q_{max} and optimum density k_{max}, respectively.

Macroscopic fundamental diagram (MFD) of urban road networks emerged as the primary modeling tool enabling aggregated modeling and control approaches for traffic management. It is insensitive to small changes in demand which makes it a perfect tool for monitoring the effects of traffic control strategies (Amini *et al.*, 2018). Based on the findings of previous research (Geroliminis and Daganzo, 2007) it was observed that an urban region with roughly homogeneous accumulation (e.g., small spatial edge density heterogeneity) can be modeled using the MFD, which provides a unimodal, low-scatter, and demand-insensitive relationship between density and trip completion flow. Recent research has proven that an MFD exists on neighborhood-sized sections of cities independently of the demand (Geroliminis and Carlos F. Daganzo, 2008).

When building a user-centric driving assistance system the ultimate goal is to improve efficiency. The efficiency of urban traffic is vital for the optimization of traffic flow, and can be achieved through the management of traffic lights, detection and management of road accidents, minimization of traffic delays, or even through improving parking services. On the other hand, efficiency from the point of view of the driver itself focuses on improving driving behavior with the goal to minimize environmental impact and fuel consumption, although these behaviors also often lead to improvements in safety and comfort (Paúl *et al.*, 2018). The great challenge here is to investigate the link between traffic and driving efficiency. In other words, is it possible to enhance traffic efficiency through the improvement of individual driving efficiency?

In this dissertation, we aim at answering this question by estimating the actual impact of applying a personalized driving recommendation system, that improves individual driving performance in terms of safety, on the road network performance. In order to achieve this goal, we take advantage of the findings outlined by (Geroliminis and Carlos F Daganzo, 2008) to first estimate MFDs before and after the application of the proposed system and then, evaluate comparatively the differences that will arise.

In order to give a direction towards addressing the aforementioned challenges, Table 2-2 presents some indicative solutions based on some studies that have tried to efficiently overcome these prevailing challenges.

Table 2-2. Existing challenges in driving behavior analysis using smartphone crowd-sensing and indicatively proposed solutions

Challenge	Proposed Solution	Selected Citations
Enhancing data representativeness, availability and quality	• Incentives • Gamification aspects • Convince the crowd for the usefulness and importance of driving behavior understanding in traffic and road safety • Utilize low-power wireless networks • Upload data when a Wi-Fi connection is available • Share sensing data among multiple systems • Change sampling rates • Anonymity, pseudonymity, spatial cloaking • Data perturbation and aggregation • Feature selection (Filter, wrapped methods) • Anomaly detection techniques	(Hossain, 2012; Christin, 2016; Etemad *et al.*, 2018; J. Wang *et al.*, 2018; Yen *et al.*, 2019)
Identify the context from the data	• Data fusion • Filtering algorithms • Feature Engineering	(Wahlström *et al.*, 2015)
Detect abnormal patterns and risky drivers	• Machine learning approaches • Determine universal thresholds for each feature	(Bejani and Ghatee, 2018)
Modeling efficiency, transfer learning and explainability	• Apply resampling techniques • Generate synthetic samples • Transfer learning • Additional features • Big data analysis instead of small experimental datasets • Outlier detection	(Yuan and Raubal, 2012; Hu *et al.*, 2018; Maldonado and López, 2018; Roy *et al.*, 2018)
Raising awareness and changing attitudes	• Incorporate UBI schemes • Gamification	(Tselentis *et al.*, 2017; Vlahogianni and Barmpounakis, 2017b)
Lack of personalization	• Build and validate driver models based on real data • Observe real driving behavior	(Hasenjager and Wersing, 2018; Ponomarev and Chernysheva, 2019)
Real-time operation	• Artificial Intelligence • Prioritize tasks • Efficient memory management	(Shukla *et al.*, 2018)
Moving from user-centric systems to network level management	• Estimate impact of user-centric solutions • Context-aware RL solutions for user-centric systems	(Geroliminis and Daganzo, 2008; Amini *et al.*, 2018)

2.6 Research questions

Based on the results of the literature review, the research questions of this doctoral dissertation are formulated as shown below together with the proposed approach for addressing them:

Question 1 (Q1)

Which are the main driving profiles that cover the wide range of driving behavior and how can they be identified by exploiting smartphone data?

Question 2 (Q2)

Is it possible to classify the overall driving behavior of drivers into groups that share common driving characteristics, and, if so, to what extent could it be classified?

Question 3 (Q3)

Could Artificial Intelligence techniques be exploited within the framework of a driving recommendation system and ensure the requires degree of personalization of the produced recommended actions?

Question 4 (Q4)

Which is the most appropriate Reinforcement Learning algorithm for supporting human decision making?

Question 5 (Q5)

Is there a link between raising self-awareness and improving conditions of the entire network? To what extent could the improvement of individual behavior affect traffic conditions?

Question 6 (Q6)

What kind of impact does the controlling of individual driving behavior have on driving and road safety?

Question 7 (Q7)

How are emissions affected by the controlling of individual driving behavior? Is there a significant change on environmental conditions when drivers improve their behavior?

3 METHODOLOGICAL APPROACH

In this section, the problem overview is described with respect to the specific objectives set within this dissertation, together with the conceptual assumptions made. Subsequently, the main structure of the methodological approach followed is presented, together with a thorough description of each methodological step. Finally, the theoretical background of the methods used in this dissertation is given in details.

3.1 Problem overview and assumptions

Driving behavior analysis has occupied researchers for many decades due to the complexity of the process of decision-making during driving, but also due to the inability of researchers to delineate driving behavior by defining thresholds that enclose specific driving habits. In this dissertation it is attempted to overcome this limitation by exploiting a large-scale naturalistic driving dataset and try to identify the underlined driving behavior patterns straight from the data. For this purpose, an unsupervised learning approach will be followed so that observed driving behavior is clustered into appropriate groups which share common driving characteristics (driving profiles). In line with previous research, here as well, it is assumed that every single driver exhibits a specific type of behavior during a trip which may be characterized by either safe or unsafe driving maneuvers. More precisely, it is assumed that every driver chooses to driver in a specific manner throughout their entire trip. One may think that a driver may change their behavior even during the same trip, fact that is not far from reality, yet due to data availability limitations, computational and operational limits researchers have not investigated the dynamics of driving behavior on such microlevel. The in-depth analysis of driving behavior at a time-window level of a trip would have been of great importance within personalized ADAS systems that operate real-time. In this dissertation, in order to design a self-determining system of providing personalized recommendations based solely on the driver himself and their driving behavior, not taking any knowledge neither about the environment nor the rest of the traffic, the approach of analyzing behavior at a trip-level is followed.

In addition, it has been previously shown that each individual's driving behavior is described by great rates of volatility meaning that the driver alters the way they drive on every trip and therefore, do not have a stable driving profile (Mantouka et al., 2018; Tselentis, 2018). Findings in (Mantouka et al., 2019) revealed that drivers behave differently every time, performing trips that fall within each one of the recognized categories of safe and unsafe driving style (trip-level driving profiles). For the purpose of this work, and taking into account the fact that specific driver profiles cannot be identified due to the high variability of behavior as described above, the estimation of the overall driving behavior of a specific driver is based on a simple statistical rule. The identified driving profiles are ranked based on a safety and cautiousness scale and then, the average behavior of each driver can be derived as the average of the incidence of each driving profile. In this way drivers can be divided into groups that indicate the rate of occurrence of safe and unsafe behaviors while driving accordingly.

Finally, in this work it is assumed that the way a driver chooses to drive is not independent from the environment they drive in, the road geometry and the traffic conditions. Previous research has shown that speed selection is highly correlated with the speed limit, road illumination and visibility, road geometry and weather conditions besides drivers' characteristics and perceptions (Sadia *et al.*, 2018; Liu *et al.*, 2020; Zolali *et al.*, 2021). Furthermore, deceleration decisions are highly dependent on each driver's risk perception as well as on other external parameters such as the existence of traffic signals and traffic lights, and the behavior of the leading car (Li *et al.*, 2020). Therefore, in order to achieve the goal of developing an independent driving recommendation system that takes into account exclusively driving behavior, it was considered that the driving style that one chooses to adopt is determined solely by the choices of acceleration. In other words, it is assumed that the only parameter of driving behavior over which the driver has absolute control and its value is determined solely by the personal preferences of each driver is acceleration. For this reason, the proposed system is trained to suggest the most appropriate alterations in acceleration choices of each driver so that to improve driving safety in both a short- and long-term. The methodological approach followed to achieve this goal is thoroughly described in the next section.

3.2 General methodological framework

Developing a personalized driving recommendation system is not a trivial task, as state-of-the-art methods are required to handle a number of challenging tasks from driving behavior identification through smartphone data to agents' efficient training for learning driving behavior's dynamics using deep learning. In order to develop the Self-Aware Driving Recommendation System, a structured procedure is followed, as shown in Figure 3-1.

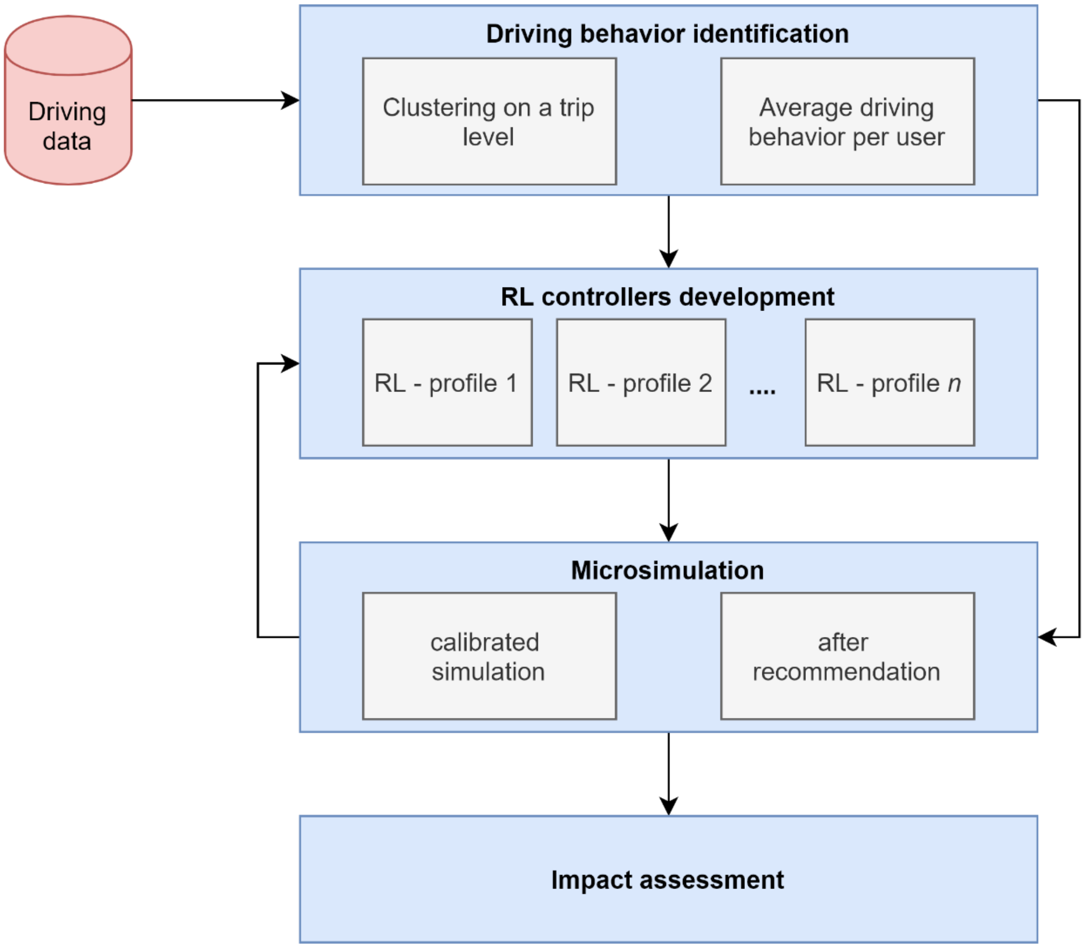


Figure 3-1. Methodological approach for the development and assessment of the Recommendation Framework

Starting from raw measurements of GPS location, acceleration and speed, as provided by a telematics application established on smartphone devices, driving features are defined that describe short-term and long-term driving behavior. Following, these features are utilized in an unsupervised learning framework to identify driving profiles that can be used to describe each driver's overall driving behavior (Q1, Q2). Driving behavior is defined at:

- a trip level, which corresponds to the way the driver performed a specific trip, and
- a user level, which corresponds to the overall driving behavior of a specific driver in all of his trips (driving footprint).

Unsupervised learning is used in exploratory analysis since it has the ability to automatically identify structure in non-annotated data, and therefore is the most appropriate technique to detect different driving behavior patterns which are obscured in the data. Unsupervised learning, namely clustering, can be either soft or hard: in hard clustering, each data point either belongs to a cluster completely or not, while in soft clustering, each data point is assigned a probability or likelihood to be in each of the clusters (Bora and Gupta, 2014). For the purpose of this research a hard clustering technique is implemented so that each driving trip is assigned to a specific driving profile which indicates that the driver has chosen to exhibit a specific driving behavior during the trip. K-means clustering is the most well-known hard clustering technique characterized by small computational requirements, robustness and interpretability.

Here, a two-level k-means clustering algorithm is implemented, in a selection of driving features, in order to distinguish aggressive from non-aggressive trips within the first level, and then further distinguish between risky and distracted driving with the second level of clustering. Several driving features' combinations were tested in the context of clustering, the performance of which was evaluated through three distance-based metrics (silhouette index, dunn's index, and the Calinski-Harabasz criterion). After this procedure, each trip was assigned to a specific driving profile (Q1), and then, using statistical measurements, the overall driving behavior of each driver is identified (Q2).

Once driving behavior per trip is identified, and all drivers were separated into groups based on their overall driving behavior, the driving recommendation framework is designed and the appropriate algorithms are developed, using state-of-the-art Reinforcement Learning models (Q3). Reinforcement Learning is considered a powerful tool due to its ability to learn optimal policies of behavior, within large and complex environments, by a continuous process of rewarding or punishing agents correspondingly to their actions (Morales, 2020). Reinforcement Learning is a data-driven method: Drivers who are found to adopt a safe driving behavior possess the knowledge of how to avoid the performance of unsafe driving maneuvers. In order to exploit this knowledge straight from the observed behavior, a reinforcement learning model is developed to learn how actions are taken under different conditions and map actions to specific states of the environment, namely to learn the optimal policy, without any previous knowledge. Speed and acceleration choices on a trip level (average and maximum values and variation) can describe the current state of a driver's behavior and their adjustment in future trips (either towards an improved or worsened behavior) can be thought of as an action; quite similarly to the definition of a reinforcement learning framework. The aim of the Reinforcement Learning algorithm is to learn the optimal policy and suggest the appropriate action that leads to the best possible behavior. Specifically, when dealing with driving behavior recommendations, every action refers to an adjustment of the vehicle's kinematic characteristics including the adjustment of the vehicle's speed and acceleration, which span within a continuous range of values. To this end, the RL algorithm developed in this work should retain one extra property, the ability to handle continuous state and action spaces (Q4).

In order to create a system that recognizes driving style and suggests improvements that match the way everyone chooses to drive, we develop separate RL controllers one for each driver type. The RL agents follow an actor-critic approach based on the Deep Deterministic Policy Gradient algorithm (Q4). Both the actor and the critic are implemented as deep artificial neural networks, the hyperparameters and the structure of which emerge after an exhaustive grid search. The algorithms are trained using sequences of driving trips of the same driver as input, while the output of each RL controller is the optimal alteration in the acceleration of each driver, given the way they drove in their previous trip.

Finally, the impact of improving individual driving behavior is assessed through a comparative before-after microsimulation analysis, with respect to road safety, traffic and the environment (Q5, Q6, Q7). Using the road network of Athens, Greece, a microsimulation scenario for the morning rush hour demand, was set. For the initial conditions of the network the vehicles move according to the characteristics governing each of the driving behaviors detected in the first step of the methodological framework, while the traffic composition is based on the actual distribution of trips over the driving profiles. In this way, driving diversity is ensured between the vehicles and the traffic conditions in the network are simulated as realistically as possible.

Thereupon, a step-wise procedure is followed in order to extract all the essential information from this first simulation scenario:

1. 10 replications of the simulation scenario are performed, all corresponding to exactly 3600 seconds of running time per replication, in order to ensure the robustness of the results and eliminate the randomness of the findings.
2. The basic traffic flow parameters, namely flow, speed and density, are estimated for the entire network and their mean values are used for the construction of the Macroscopic Fundamental Diagrams, which are considered as the KPI of traffic conditions (Q5).
3. Using the Surrogate Safety Assessment Model (FHA, 2008) critical conflicts are estimated by exploiting the information regarding the trajectories followed by each vehicle in the simulation. The number of critical conflicts is estimated in an aggregated-level as well as per vehicle, and acts as a KPI for the safety of the network (Q6).
4. The amount of harmful air pollutants is estimated by category of pollutant, in an aggregated level and per vehicle as well. The number of emissions is considered as the KPI of environmental conditions (Q7).
5. Statistical characteristics (mean, max, min, quartiles) of all driving parameters are calculated, including acceleration, deceleration, speed, and speeding, for each vehicle.
6. The driving state of each vehicle is used as input to the corresponding RL controller, which produces the optimal alteration of the acceleration.
7. For each individual vehicle the optimal acceleration is calculated with correspondence to the action proposed by the recommendation system. A table containing the acceleration values per vehicle is used as input for the second run of simulation.

In the second run of the simulation, all the conditions remain the same as before, namely the simulation time, the demand, and the routes followed by each vehicle, with only exception the acceleration values of the vehicles. In this case, the recommendation is followed, and after 10 replications, the crucial KPIs of traffic, safety, and environmental conditions are estimated. To end, a critical discussion on the differences found in network's performance is provided and the impact of applying the personalized recommendations is quantified.

3.3 Theoretical Background

The problem of identifying driving profiles from real-world driving data is generally an unsupervised learning problem. In this dissertation, the most popular and at the same time the most robust method for grouping unlabeled data, k-means clustering is used. The basic principles of clustering along with its goodness-of-fit measures are presented in this section.

As far as it concerns the recommendation system, its development is based on the idea of an agent that can learn how individuals choose to drive and how their behavior evolves between sequential trips. By learning this information, the agent should then be able to learn which is the most likely transition to safer behavior, given the current way of driving. This learning process is approached in the modern literature by reinforcement learning algorithms. Later in this section, the fundamentals of Reinforcement Learning and the most popular algorithms are presented with a special focus on the algorithm that matches the objectives of this dissertation, the Deep Deterministic Policy Gradient.

Finally, for the sake of completeness of the theoretical background, reference is made to traffic flow theory fundamentals that describe traffic conditions on a road network, yet without further analysis since these are the most basic knowledge in the field of transport engineering.

3.3.1 Clustering

Clustering is a well-known task of dividing a set of observations into a number of groups so that the observations within the same group are similar. The most widely used clustering technique is K-means clustering, where a cluster can be thought as a group of data points whose interpoint distances (intra-cluster similarity) are small compared with the distances of points outside of the cluster (inter-cluster similarity) (Kanungo *et al.*, 2002). Specifically, the following definitions have been given for the distance between the objects of different clusters and the objects of the same clusters Figure 3-2:

- Inter-cluster distance is the distance between two objects belonging to two different clusters
- Intra-cluster distance is the distance between two objects belonging to the same cluster

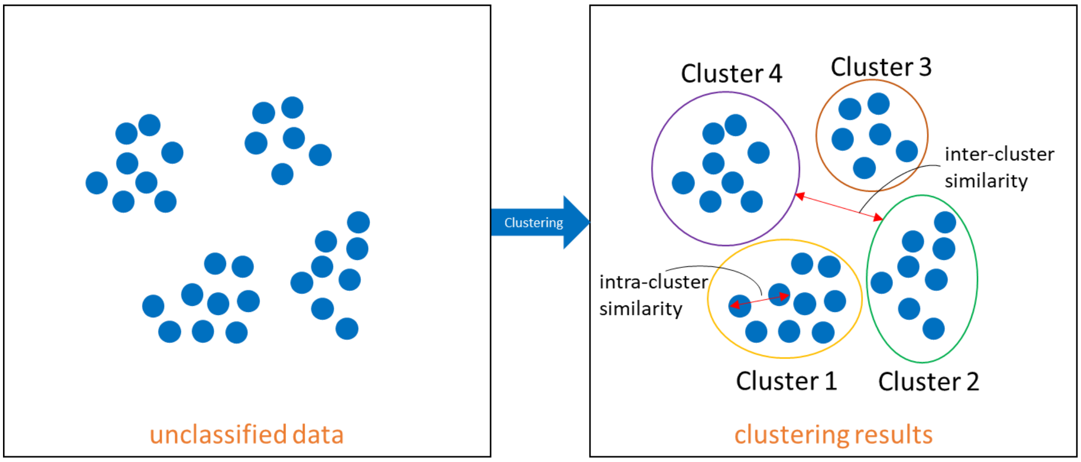


Figure 3-2. Clustering results and similarity measures

For each data point X_n, a corresponding set of binary indicator variables $r_{nk} \in \{0.1\}$ are introduced, where $k = 1, \ldots\ldots, K$ describing which of K clusters the data point X_n is assigned to, so that if a data point is assigned to cluster k then $r_{nk} = 1$, and $r_{nj} = 0$ for $j \neq k$. Then, an objective function is defined, given by:

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \| x_n - \mu_k \|^2 \tag{1}$$

Which represents the sum of squares of the distances of each data point to its assigned vector μ_k, where μ_k represents the center of the k^{th} cluster. The goal is to find values for $\{r_{nk}\}$ and the $\{\mu_k\}$ so as to minimize J (Bishop, 2006). The Pseudo-code of the Lloyd's K-Means algorithm is the following (Mohd *et al.*, 2012):

Algorithm 1: Lloyd's K-Means Algorithm

Input Replay memory size M, batch size d, number of episodes E, and number of time steps T
Initialize Main network weights θ
Initialize Target network weights θ^-
Initialize Replay memory
for $e = 1, \ldots, E$ **do**
  Initialize state s_1, and action a_1
  for $t = 1, \ldots, T$ **do**
    Take action $a_t = argmax_a Q^{\pi}(s_t, a; \theta)$ with probability $1-\epsilon$ or a random action with probability ϵ
    Get reward r_t and observe next state s_{t+1}
    if Replay capacity M is full then
      Delete the oldest tuple in memory
    end if
    Store the tuple (s_t, a_t, r_t, s_{t+1}) to replay memory
    Sample random d tuples from replay memory

$$y_t = \begin{cases} r_t, & if\ t = T \\ r_t + \gamma \max_a Q^{\pi}(s_{t+1}, a_{t+1};\ \theta_t^-), & otherwise \end{cases}$$

    Perform policy gradient using y_t for updating θ
    Update target network every N step, $\theta^- = \theta$
  end for
end for

Clustering validation measures evaluate the goodness of clustering results, and are considered as key for the success of clustering applications. There is a variety of validation measures, such as Silhouette index, Dunn's index, Davies-Bouldin index, R-square, Hubert's Γ statistic, Calinski-Harabasz criterion and many more. For the purposes of this dissertation, the following three measures are used (Maulik and Bandyopadhyay, 2002; Kraft, 2012):

Silhouette index

For a given cluster, X_j *(j=1,….,c)*, this method assigns to each sample of X_j a quality measure, *s(i) (i=1, ….,m)*, known as the Silhouette width, which is defined as:

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \qquad (2)$$

where *a(i)* is the average distance between the i^{th} sample and all of the samples included in X_j; 'max' is the maximum operator, and *b(i)* is the minimum average distance between the i^{th} sample and all of the samples clustered in X_k (k =1,….,c; k ≠ j). From this formula, it follows that $-1 \leq s(i) \leq 1$. When a *s(i)* is close to 1, one may infer that the i^{th} sample has been assigned to an appropriate cluster. When a *s(i)* is close to 0, it suggests that the i^{th} sample could also be assigned to the nearest neighboring cluster. If *s(i)* is close to −1, one may argue that such a sample has been assigned to a wrong cluster. Thus, for a given cluster it is possible to calculate a cluster Silhouette index S_j, which characterizes the heterogeneity and isolation properties of such a cluster:

$$Sj = \frac{1}{m}\sum_{i=1}^{m} s(i) \qquad (3)$$

where *m* is the number of samples in S_j.

Dunn's index

For any partition $U \leftrightarrow X : X_1 \cup ... X_i \cup ... X_c$ where X_i represents the i^{th} cluster, the Dunn's index is defined as follows:

$$D(U) = \min_{1 \leq i \leq c} \left\{ \min_{\substack{1 \leq j \leq c \\ j \neq i}} \left\{ \frac{\delta(X_i, X_j)}{\max_{1 \leq k \leq c} \{\Delta(X_k)\}} \right\} \right\} \qquad (4)$$

where *δ*(X_i, X_j) defines the inter-cluster distance, the distance between clusters X_i and X_j, *Δ*(X_k) represents the intra-cluster distance of cluster X_k , and *c* is the number of clusters of partition *U*. This index is easy to implement and has a low computational complexity. It is obvious that large values of *D(U)* indicate better clustering.

Calinski-Harabasz criterion

The cluster index of Calinski-Harabasz is calculated using the following equation:

$$CH(K) = \frac{[traceB/K - 1]}{[traceW/N - K]} \qquad for \qquad K \in \mathbb{N} \tag{5}$$

where B denotes the inter-cluster error, the error sum of squares between different clusters:

$$traceB = \sum_{k=1}^{K} |C_k| \left\|\overline{C_k} - \overline{x}\right\|^2 \tag{6}$$

and W the squared differences of all objects in a cluster from their respective cluster center (intra-cluster):

$$traceW = \sum_{k=1}^{K} \sum_{i=1}^{N} w_{k,i} \left\|x_i - \overline{C_k}\right\|^2 \tag{7}$$

The maximal achieved value of the Calinski-Harabasz criterion indicates the best clustering of the data.

3.3.2 Reinforcement Learning

Reinforcement learning has potential in the area of intelligent transportation due to its generality and ability to achieve human level performance in many complex tasks. This approach to learning is inspired by behaviorist psychology, where human and animal behavior is studied from a reward and punishment perspective. In the structure of a RL system, an agent interacts with an environment. After every discrete time t, the agent implements an action a, the environment changes from the previous state s_t to s_{t+1}, and the agent gets a corresponding reward r_t (Figure 3-3).

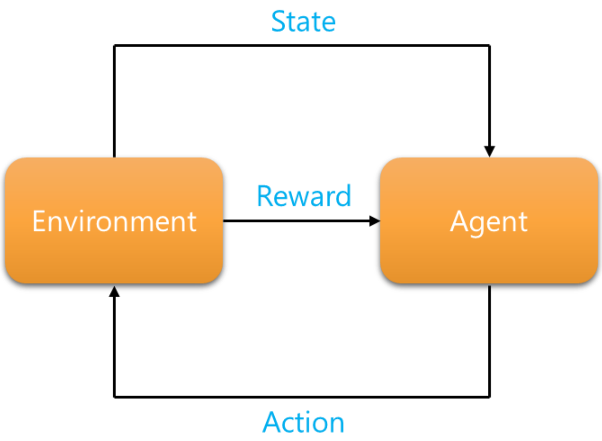


Figure 3-3. Main Reinforcement Learning components representation

The *environment* is represented by a set of variables related to each problem. The combination of all the possible values this set of variables can take is referred to as the *state space*. A *state* is a specific set of values the variables take at any given time. Agents may or may not have access to the actual environment's state; in any case, agents can observe something from the environment. The set of variables the agent perceives at any given time is called an *observation*. The combination of all possible values these variables can take is the *observation space*. (Nowé and Brys, 2016)

In most of RL problems the agent has access to all the information that describe the environment and therefore state and observation terms are used interchangeably. At every state, the environment makes available a set of *actions* the agent can choose from. The set of all actions in all states is referred to as the *action space*. The environment changes states as a response to the agent's action following the so-called *transition function*. After a transition, the environment emits a new state and also provide a *reward* signal as a response. The function responsible for this mapping is called the *reward function* (Morales, 2020):

$$r(s,a) = \mathbb{E}[R_t | S_{t-1} = s, A_{t-1} = a] \qquad (8)$$

Usually, RL problems are modeled as a discrete-time Markov Decision Process (MDP) with a tuple of $(S, A, P_{ss'}, R, \gamma)$, which includes a state space S; an action space A of all possible actions; a transition function $P(s_{t+1}|s_t,\ a_t)$ which measures the probability of obtaining the next state s_{t+1}given a current state-action pair (s_t, a_t); the immediate reward $R(s_t, a_t)$ achieved at each state-action pair, and $\gamma\ \in (0,1)$ that denotes a discount factor (Tuyen and Chung, 2017).

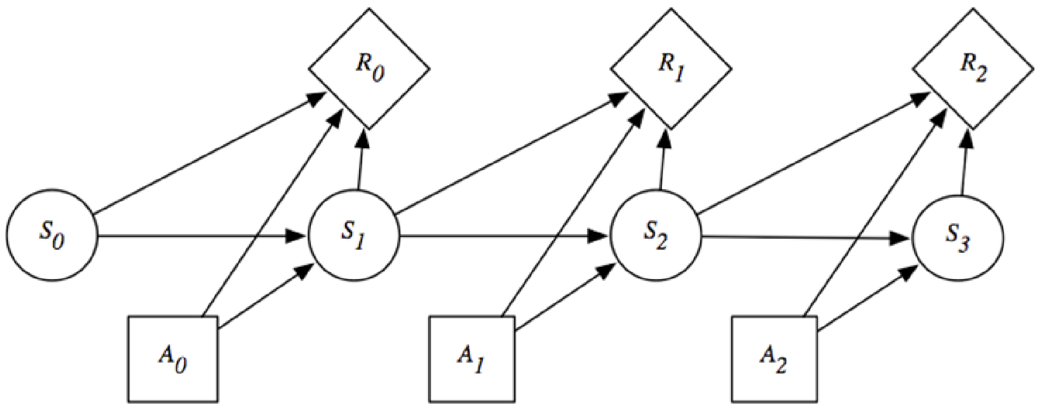


Figure 3-4. Markov Decision Process

It is called a Markov Decision Process, since the state signal is assumed to have the Markov property:

A stochastic process has the Markov property if the conditional probability distribution of future states of the process (conditional on both past and present states) depends only upon the present state, not on the sequence of events that preceded it (Bhattacharya and Waymire, 2007).

A sequence of state-action pairs (s_t, a_t) creates a trajectory ξ_t, also referred to as *episode*, with discounted cumulative reward given by:

$$R(\xi) = \sum_{t=0}^{T} \gamma^t r(s_t, a_t) \tag{9}$$

The algorithm used by the agent to determine its actions, i.e., its behavior, is commonly referred to as a *policy*, denoted as *π*. A policy is a function that prescribes actions to take for a given state. In some cases, the policy may be a simple function or lookup table, whereas in others it may involve extensive computation such as a search process (Sutton and Barto, 2018). The policy is the core of a reinforcement learning agent in the sense that it alone is sufficient to determine behavior. In general, policies may be either deterministic or stochastic, specifying probabilities for each action.

If the algorithm estimates the optimal policy without using or estimating the dynamics (transition and reward functions) of the environment, then it is called a *model-free* algorithm. Otherwise, if the algorithm uses the transition function (and the reward function) in order to estimate the optimal policy then it is referred to as *model-based* algorithm.

3.3.2.1 Q-learning

Q-learning is a model-free reinforcement learning algorithm, based on the well-known Bellman Equation. When the agent follows a certain policy *(π)*, then the Value (v_π) of a particular state is determined by the immediate reward plus the value of successor states emerged from the Bellman Expectation Equation (Van Otterlo and Wiering, 2012):

$$v_\pi(s) = \mathbb{E}[R_{t+1} + \gamma v_\pi(S_{t+1}) | S_t = s] \tag{10}$$

Where $\mathbb{E}$ stands for expectation and *γ* is the discount factor. Re-writing the above equation in the form of the Q-value results in the following:

$$Q^\pi(s,a) = \mathbb{E}[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots \quad | s, \alpha] = \mathbb{E}s'[r + \gamma Q^\pi(s', a') | s, a] \tag{11}$$

In other words, Q-value (*Q*) is similar to Value, except that it takes as an additional parameter, the current action (*a*). $Q^\pi(s,a)$ refers to the long-term return of the current state *s*, taking action *a* under policy *π* (Van Otterlo and Wiering, 2012). The data structure used to match the best action at each state based on the Q-values, is refer to as Q-table. The algorithm of Q-learning is schematically presented in Figure 3-5.

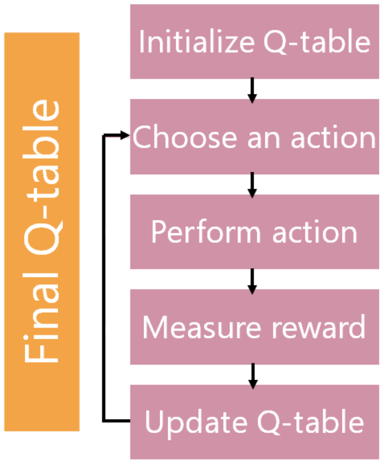


Figure 3-5. Schematic representation of Q-learning algorithm

Despite the fact that Q-learning has revolutionized the capabilities of reinforcement learning, it fails to perform in big and complex environments with thousands of states and actions pairs. In order to overpass such limitations researchers approximated Q-values with machine learning models such as neural networks.

3.3.2.2 Deep Q Network

Advances in deep learning allowed significant progress in Reinforcement Learning with the introduction of Deep Q Network (DQN) algorithm which is capable of handling discrete but low-dimensional action spaces (Mnih *et al.*, 2013; Lillicrap *et al.*, 2016). In a DQN the state is given as the input and the Q-value of all possible actions is generated as the output. The comparison between Q-learning and DQN is perfectly illustrated in Figure 3-6.

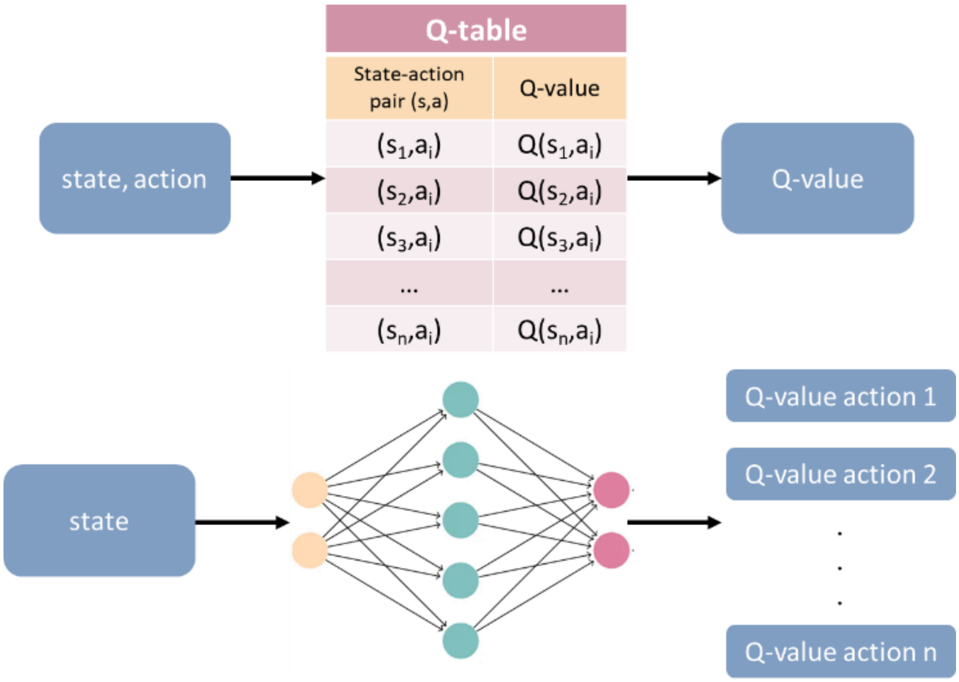


Figure 3-6. Comparison between Q-learning & deep Q-learning

Two techniques are essential in order to train a DQN; Experience Replay and a Target Network.

The basic idea behind *experience replay* is to storing past experiences of the agent and then, using a random subset of these experiences to update the Q-network, instead of using the most recent experience. This is done in order to prevent action values from oscillating or diverging tragically. The replay memory contains a collection of experience tuples in the form of (s_t, a_t, r_t, s_{t+1}). The act of sampling a small batch of tuples from the replay memory is referred to as experience replay.

The *target network* is basically a copy of the Q-network which allows more stable training. Specifically, when an update is performed to the Q-networks' parameters in order to make *Q(s, a)* closer to the desired result, it consequently alters the value produced for *Q(s', a')* as well. This can make the training very unstable and thus, to overcome this issue, a target network with the same structure as the Q-network is used. It should be highlighted that the target network's parameters are not trained, but they are periodically synchronized with the parameters of the main Q-network.

The pseudocode of the DQN Algorithm is as shown below:

Algorithm 2: DQN Algorithm

Input Replay memory size M, batch size d, number of episodes E, and number of time steps T
Initialize Main network weights θ
Initialize Target network weights θ^-
Initialize Replay memory
for $e = 1, \ldots, E$ **do**
  Initialize state s_1, and action a_1
  for $t = 1, \ldots, T$ **do**
    Take action $a_t = argmax_a Q^{\pi}(s_t, a; \theta)$ with probability $1-\epsilon$ or a random action with probability ϵ
    Get reward r_t and observe next state s_{t+1}
    if Replay capacity M is full then
      Delete the oldest tuple in memory
    end if
    Store the tuple (s_t, a_t, r_t, s_{t+1}) to replay memory
    Sample random d tuples from replay memory

$$y_t = \begin{cases} r_t, & if\ t = T \\ r_t + \gamma \max_a Q^{\pi}(s_{t+1}, a_{t+1};\ \theta_t^-), & otherwise \end{cases}$$

    Perform policy gradient using y_t for updating θ
    Update target network every N step, $\theta^- = \theta$
  end for
end for

3.3.2.3 Policy Gradient

Despite the fact that DQN achieved huge success in higher dimensional problems, the action space remains discrete. However, in a great variety of tasks, especially physical control tasks, the action space is continuous. In order to solve more complex problems, where both state and action spaces are continuous, Lillicrap *et al.* (2016) took advantage of the Deterministic Policy Gradient algorithm (Silver *et al.*, 2014), and introduced the Deep Deterministic Policy Gradient algorithm (DDPG), which is a model-free, off-policy actor-critic algorithm which uses deep function approximators that can learn policies in high-dimensional, continuous action spaces. The basic idea behind the DDPG algorithm is that it follows a stochastic approach for exploring all possible actions but estimates a deterministic policy.

A key feature of the DDPG algorithm is its simplicity since it requires only a straightforward actor-critic architecture and a learning algorithm, making it easy to implement and scale up to a series of difficult problems (Lillicrap *et al.*, 2016). In the actor-critic algorithm, the **actor**, namely the policy function, generates an optimal action given the current state. In other words, an actor is used to tune the parameter $\boldsymbol{\theta}$ for the policy function, i.e., decide the best action for a given state. A **critic** is used to evaluate the policy function estimated by the actor based on the temporal difference (TD) error.

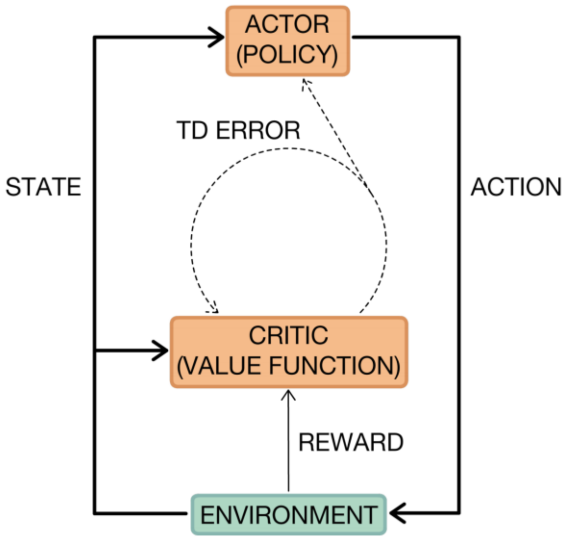


Figure 3-7. Schematic representation of the Actor-Critic algorithm

More specifically, the actor holds the policy function $\mu(s|\theta^{\mu})$ which generates an action given the current state. The critic evaluates an action-value function $Q(s,a|\theta^{Q})$ based on the output from the actor, as well as the current state (P. Wang *et al.*, 2019) to ensure that it is the optimal. The actor and critic are designed with neural networks. The deterministic policy gradient theorem (Silver *et al.*, 2014) provides the update rule for the actor network.

Policy Gradient Theorem: The derivative of the expected reward is the expectation of the product of the reward and gradient of the log of the policy π_{θ}.

The critic network is updated from the gradients obtained from temporal-difference errors, which can be formulated as follows (P. Wang *et al.*, 2019):

$$\nabla_{\theta_\mu}\mu \approx E_\mu[\nabla_a Q(s,a|\theta^Q)|_{s=s_t,a=\mu(s_t)}\ \nabla_{\theta_\mu}\mu(s|\theta^\mu)|_{s=s_t}] \quad (12)$$

For a mini-batch, the critic network is updated by minimizing the loss in *(4)*. The actor policy is updated with sampled policy gradients as shown in *(5)*.

$$L = \frac{1}{N}\sum_i \left(y_i - Q(s_i,a_i|\theta^Q)\right)^2, where\ y_i = r_i + \gamma Q'(s_{i+1},\mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'} \quad (13)$$

The updates of the target critic and target actor networks are as follows:

$$\begin{cases}\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}\end{cases} \quad (14)$$

where τ is an update parameter and can be set as $\tau \ll 1$.

The pseudocode for implementing the DDPG algorithm is as follows (Lillicrap *et al.*, 2016).

Algorithm 3: DDPG Algorithm

Randomly initialize critic $Q(s,a|\theta^Q)$ and actor $\mu(s|\theta^\mu)$ with weights $\theta^Q$and θ^μ
Initialize target network Q' and μ' with weights $\theta^{Q'} \leftarrow \theta^Q$, $\theta^{\mu'} \leftarrow \theta^\mu$
Initialize replay buffer R
for episode=1, M **do**
  Initialize a random process $\mathcal{N}$ for action exploration
  Receive initial observation state s_1
  for t=1, T **do**
    Select action $a_t = \mu(s_t|\theta_\mu) + \mathcal{N}_t$ according to the current policy and exploration noise
    Execute action a_t and observe reward r_t and observe new state s_{t+1}
    Store transition (s_t,a_t,r_t,s_{t+1}) in R
    Sample a random minibatch of N transitions (s_i,a_i,r_i,s_{i+1}) from R
    Set $y_i = r_i + \gamma Q'(s_{i+1},\mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$
    Update critic by minimizing the loss: $L = \frac{1}{N}\sum_i\left(y_i - Q(s_i,a_i|\theta^Q)\right)^2$
    Update the actor policy using the sampled policy gradient:
$$\nabla_{\theta_\mu}J \approx \frac{1}{N}\sum_t [\nabla_a Q(s,a|\theta^Q)|_{s=s_t,a=\mu(s_t)}\ \nabla_{\theta_\mu}\mu(s|\theta^\mu)|_{s=s_t}]$$
    Update the target networks:
$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$$
$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$$
  end for
end for

Contrary to the value-based methods, in policy-gradient methods the aim is to maximize a performance objective. In value-based methods, the main focus is to learn to evaluate policies. For this, the objective is to minimize a loss between predicted and target values. More specifically, the goal is to match the true action-value function of a given policy, and therefore, to parametrize a value function and minimize the mean squared error between predicted and target values. When true target values are not available actual returns are used instead in Monte Carlo methods or predicted returns in bootstrapping methods (Morales, 2020). In policy-based methods, on the other hand, the objective is to maximize the performance of a parameterized policy, and therefore gradient ascent is used (or executing regular gradient descent on the negative performance). It's rather evident that the performance of an agent is the expected total discounted reward from the initial state, which is the same thing as the expected state-value function from all initial states of a given policy.

The main advantage of learning parameterized policies is that policies can be any learnable function. In value-based methods, one can work only with discrete action spaces, mostly because the maximum value should be calculated over all the actions. In high-dimensional action spaces, this max could be prohibitively expensive. Moreover, in the case of continuous action spaces, value-based methods are severely limited. Policy-based methods, on the other hand, can more easily learn stochastic policies, which in turn has multiple additional advantages. First, learning stochastic policies means better performance under partially observable environments. The intuition is that because arbitrary probabilities of actions can be learnt, the agent is less dependent on the Markov assumption (Morales, 2020).

3.3.2.4 Some evidence from relevant work

Over the last couple of years there has been a big increase in the application of Reinforcement Learning in a variety of problems that mostly include video games, robotics (Kober *et al.*, 2013) and autonomous vehicles (Wang *et al.*, 2018). The application of RL to solve control problems falls into two categories regarding the action space. One is discrete action space where a set of discretized actions are stored as control commands, as opposed to the continuous action space in which any real number can be sampled within the allowed threshold. In the case of discrete action spaces, the difficulty of finding an optimal action in a finite space is minimized, but the control accuracy is sacrificed.

A Q-learning method is adopted in (You *et al.*, 2018) where an autonomous vehicle learns to accelerate, brake, overtake and make turns under different highway driving scenarios taking onto account road geometry conditions. Mukadam *et al.* (2017) proposed a framework that uses deep reinforcement learning solely to obtain a high-level policy for tactical decision making for self-driving cars during lane change. In (Wang *et al.*, 2019), the authors apply DDPG for learning driving behaviors of autonomous vehicles, and particularly in the challenging task of lane change. Importantly, they do not leverage any prior knowledge of the environment and vehicle kinematics but instead, they trained the RL agent through a well-designed reward function.

There are some limited studies on the driving control field that treated the action space as discrete in order to simplify the problem or improve learning efficiency. Zhang *et al.* (2018) used Double Q-learning to learn vehicle speed control where the agent learns to either accelerate, decelerate, or maintain. An Asynchronous Actor Critic (A3C) method is applied in

(Jaritz *et al.*, 2018) to learn car control in rally games in an end-to-end framework in which the control commands were broke into 32 distinct classes. The Deep Q-network (DQN) approach is proposed in (Hoel *et al.*, 2018) to address both speed control and decision making for lane change situations.

Some studies have attempted to treat the vehicle control problem in continuous spaces. For example, P. Wang *et al.* (2018) take advantage of prior knowledge of the vehicle control mechanism and proposed a quadratic Q-function approximator to find optimal actions. Sallab *et al.* (2017) have used a deep Q- network and Deep Deterministic Actor Critic (DDAC) model in order to compare the effect of using discrete action space and continuous action space respectively, for the lane keeping task. Their findings revealed that both methods could achieve successful lane keeping behavior, however, the DDAC model showed better performance with smoothed actions. Liang *et al.* (2018) followed a Controllable Imitative Reinforcement Learning (CIRL) approach to constrict the action exploration in a controllable action space and developed a DDPG model. Kaushik *et al.* (2018) also used DDPG to learn overtaking maneuvers in continuous action space. The agent at first learns simple tasks (e.g., lane keeping) and then, moved on to complex tasks namely overtaking. Another study proposed a deep reinforcement learning approach where an RL agent learns human-like driving behavior through trial and error interactions based on a reward function that signals how much the agent deviates from the empirical data (Zhang *et al.*, 2018). Through these interactions, an optimal policy can be learned, namely the car-following model that maps in a human-like way from speed, relative speed between a lead and following vehicle, and inter-vehicle spacing to acceleration of a following vehicle is finally obtained.

In most of these studies the reward function is constructed in such a simplify way that it cannot reflect the goodness of a specific action and therefore, the agent cannot be trained towards the best possible policy. To allow efficient and effective learning, a good reward function should provide informative signals for every state-action pair and reveal useful information in the learning procedure (P. Wang *et al.*, 2019). Although it is not hard to design a reward function which gives evaluation values, it is very difficult to create an instructive reward function which can guide the RL agent to move toward the right policy.

It is evident that little work has been done in the field of applying RL algorithms for human driving guidance and control and even more less in more complex continuous environments. Therefore, there is a huge active field of research that needs to be explored and it can revolutionize design driving control systems, ADAS, cooperative Intelligent Transportation Systems (c-ITS) or even autonomous driving.

4 IMPLEMENTATION

In this section, analysis conducted in each methodological step is presented step by step, a complete view on the available data which describe naturalistic driving behavior dynamics and emerged from smartphone crowd-sensing are described and the main elements of the implementation of the methods used are discussed in details. The detailed workflow of the tasks from driving behavior identification using smartphone data to the development of the RL controllers and the impact assessment is shown in Figure 4-1. Initially, a two-level k-means clustering approach is followed to identify a variety of driving behavior patterns that emerge during a trip. The identified driving profiles include the entire range of driving preferences and habits that a driver exhibits, spanning from a cautious and safe behavior, to aggressiveness and unsafe driving habits such as distraction and risk taking. Once driving behavior on a trip level is detected, the overall driving profile of each user is estimated as the average behavior resulting from all their trips, which from this point onwards will be referred to as "user profile". Taking into consideration statistical characteristics of the user profiles, drivers were separated into to two main groups, those who adopt a typical behavior, where most of their trips are characterized by safe driving, and the rest of the drivers belong to the unsafe drivers' group since their trips are characterized by aggressiveness, distraction, riskiness or even a combination of these driving habits.

Subsequently, two driving controllers are developed based on a Reinforcement Learning algorithm; the first one produces driving recommendations for the typical drivers, while the other one produces recommendations for the unsafe drivers' group. The aim of the recommendation is to propose the most appropriate alteration in the behavior of each driver so that they improve the way they drive in terms of safety and aggressiveness extenuation. In order for these recommendations to represent real driving behaviors and not an ideal but yet unrealistic behavior, the controllers are exploiting naturalistic driving data for their training. The design of the system is based on the idea that a driver is being monitored over an adequate number of trips so that their driving behavior can be identified both on a trip and a user level. Then, driving behavior alterations are suggested that correspond to the way that each driver should drive over his next trip based on the way he drove during his last trip. The proposed system is independent to the environment where the trip is performed while focusing only on the dynamics of driving behavior of a specific driver during their trip, and therefore, the RL approach is not context-aware. The Deep Deterministic Policy Gradient algorithm is implemented since both the state and the action space are continuous as will be explained later in this section.

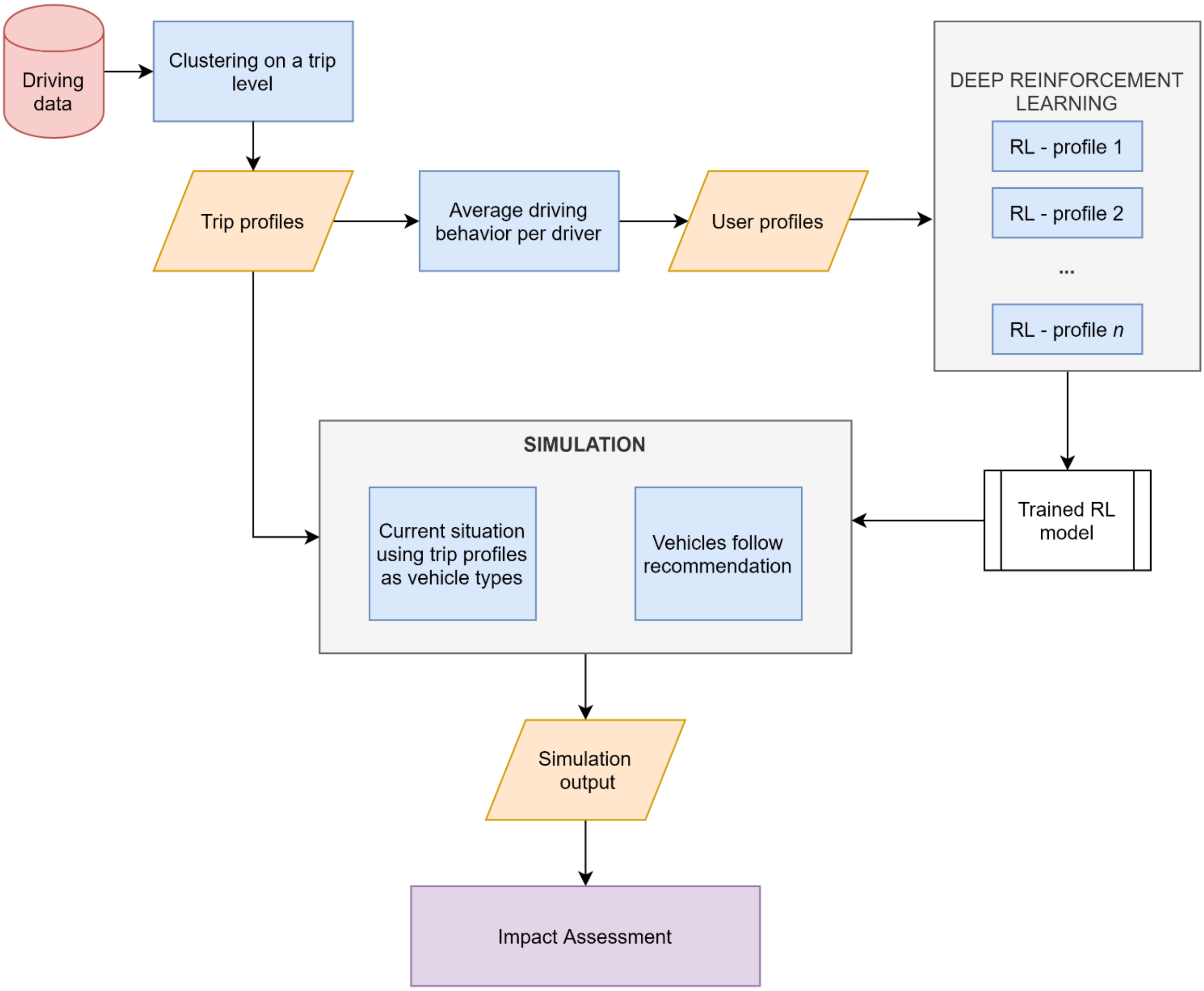


Figure 4-1. Flowchart from driving behavior identification to driving controller development and impact assessment through simulation

Once the two controllers are trained and ready to produce the most appropriate actions to each individual driver, a microsimulation scenario is set up to evaluate the impact that the application of a personalized driving recommendation system would have on traffic, safety and the environment. Two circles of simulation are performed corresponding to the initial state of the network and the performance of the network after the recommendations, respectively. The simulation scenario setting utilizes the Athens' Road network and the already calibrated transportation demand. Finally, in order to quantify the impact of the proposed system several Key Performance Indicators (KPIs) are used. The Macroscopic Fundamental Diagrams of each cycle of simulation are estimated and present the differences that emerged between the main macroscopic flow variables, flow, speed and density. As far as it concerns road safety, the number of conflicts is estimated before and after the recommendation, with a special focus on rear-end conflicts which mostly represent the impact of car-following behavior. Finally, with respect to the environmental impact, the amount of harmful air pollutants per vehicle is estimated.

The application of the proposed methodology required the development of numerous targeted scripts in both Python and R programming language. The implementation of the driving behavior identification methodology was mostly performed utilizing R programming language, while the Reinforcement Learning algorithms were developed in Python programming language, using various widespread packages such as Tensorflow, Numpy and Pandas. Additional details for the algorithmic implementation are given in the corresponding sections that follow. Moreover, several open-source software and tools were used to accomplish the objectives of this dissertation, and specifically, QGIS was used for the intuition and visualization of the data, SUMO (Simulation of Urban MObility) simulator was used to perform the simulations, large-scale data handling and xml files' modification was performed through the Atom software and the Surrogate Safety Assessment Model (SSAM) was utilized for the estimation of crucial conflicts between vehicles. A graphical demonstration of the programming language and the software exploited in each step of the methodological framework is nicely offered in Figure 4-2.

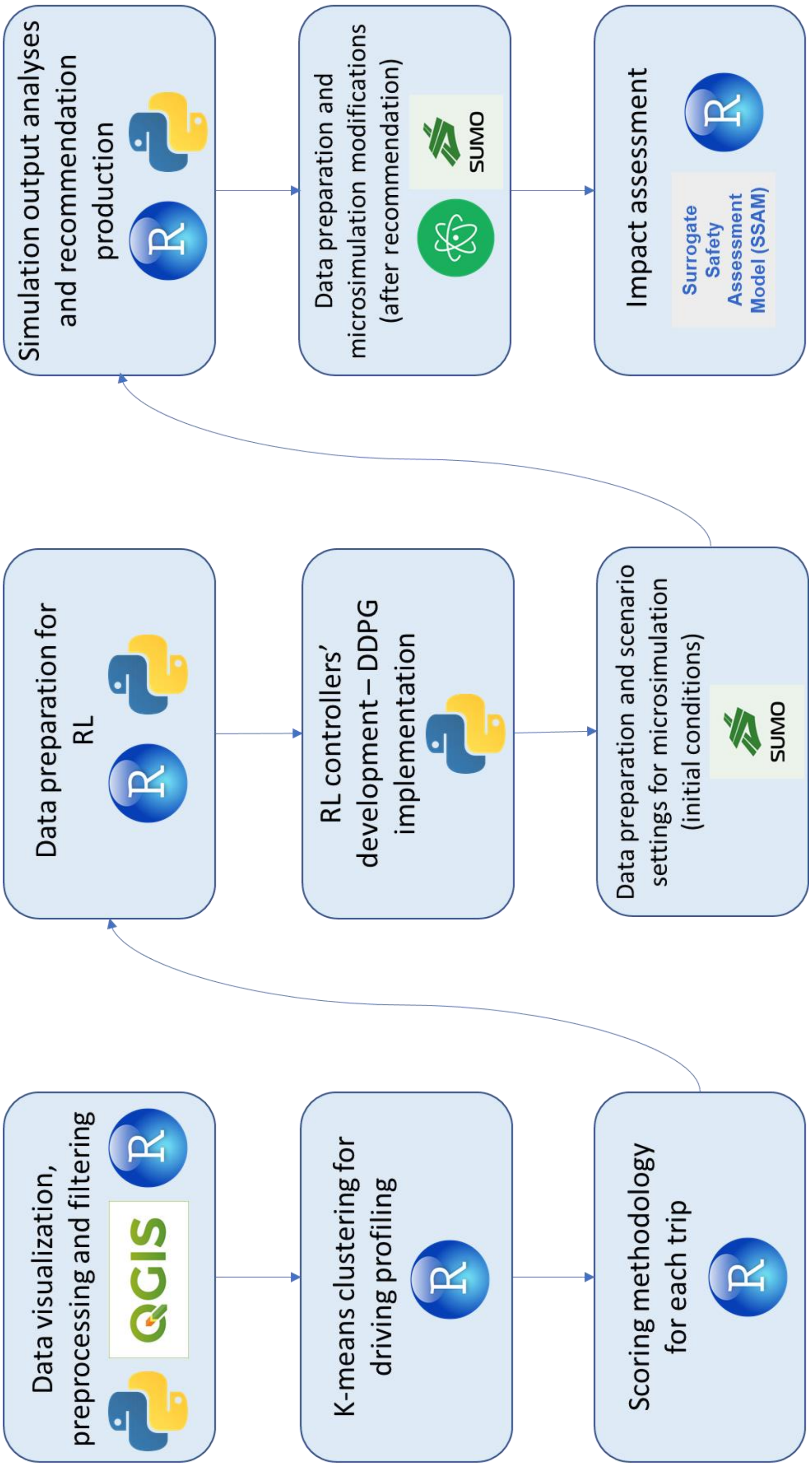


Figure 4-2. Software and tools used in each methodological step.

4.1 Naturalistic Driving Data

One of the main innovation aspects of this dissertation is the use of a large-scale naturalistic driving dataset which comes in contrast to traditional data sources such as driving simulations and on-road guided experiments. The corresponding data were collected from the OSeven application through smartphone crowd-sensing.

4.1.1 The Oseven platform

OSeven Telematics is a company that works in the fields of insurance telematics and driving behavior analysis. Since 2015, OSeven has been developing a complete system for the recording, evaluation, storage, and visualization of driving data, enabled by machine learning algorithms, driving scoring models and gamification schemes. The data recording is carried out through the OSeven smartphone application for both iOS and Android operational systems. The application exploits smartphone's embedded sensors in order to collect valuable data concerning among others trip characteristics, driving behavior, eco behavior and searching for parking behavior.

The application which is always running in the background of the smartphone's operating system, starts data recording at the beginning of every trip without requiring any user action. According to the OSeven algorithms, a trip is defined as the time period from the beginning of the driving task until a stop of driving of at least five minutes is detected. Data recording is conducted with a frequency of at least 1 Hz. Data are stored locally in each user's device, until it is wirelessly transmitted to the OSeven backend office through WiFi or mobile network data (3G/4G), based on user's preferred settings. All data stored in the OSeven backend system use advanced encryption and data security techniques in compliance with the national laws and EU directives for the protection of personal data (e.g., GDPR). The API used supports user authentication and encryption to prevent unauthorized data access. All data provided by Oseven are in a fully anonymized format. The data flow system of the OSeven platform is illustrated in Figure 4-3.

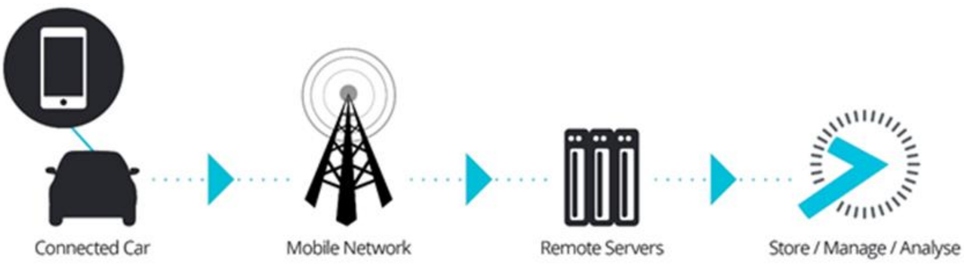


Figure 4-3. Data flow on the Oseven platform

Using a variety of criteria, the application starts to collect raw data from smartphone using accelerometer, gyroscope and GPS sensors. The accelerometer can record a smartphone's acceleration in m/s^2 in respect to gravity acceleration while the gyroscope records smartphone's angular velocity in rad/sec. Finally, GPS data are collected to record the speed of the vehicle and the coordinates of the vehicle. Since the application is using cloud-based services, after the automatic detection of the end of the trip, data are uploaded to the server for storage in an anonymized way and are ready for further process. First, data noise is excluded from the database using sophisticated data cleaning procedures, so that to correct the raw accelerometer data and specifically, to reorient the smartphone-referenced coordinate system in relation to the vehicle coordinate system (Vlahogianni and Barmpounakis, 2017a). This problem of identifying and correcting the positioning of a smartphone is treated by applying a dynamically updated reorientation algorithm which corrects the sensors' signals to address the uncertainties that stem from the arbitrary positioning of smartphones inside vehicles based on the Euler's rotation theorem (Vlahogianni and Barmpounakis, 2017a).

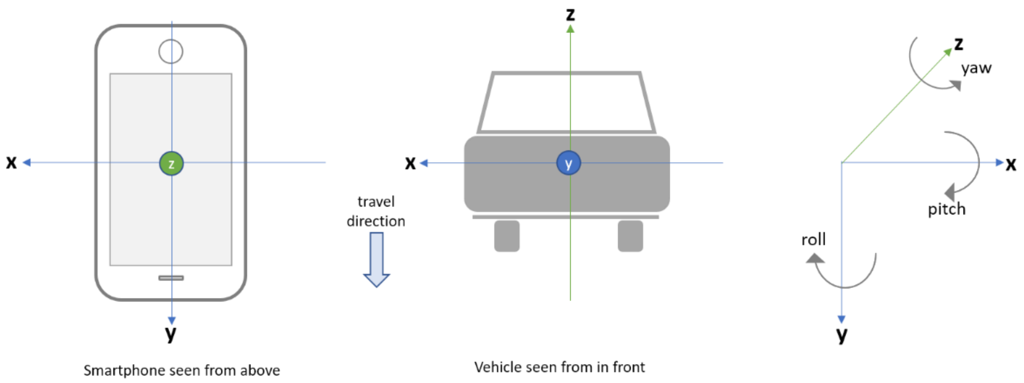


Figure 4-4. Illustration of smartphone and vehicle coordinate systems and smartphone angular rotations

Subsequently, data is converted into meaningful behavior- and safety-related parameters. Driving data processing results in the estimation of more than 400 metadata which include trip characteristics, driving behavior characteristics, parking characteristics and some more. These parameters along with additional data from external sources (e.g., maps) are subsequently exploited to implement individual driver's statistics, on all different road types (urban, highway, etc.) and under various driving conditions, enabling the creation of a large database of driving characteristics.

4.1.1.1 Some remarks on harsh events detection

One of the most critical behavior which can be detected from driving data is aggressiveness. Within this dissertation, as well as in previous research that exploits data collected through the Oseven application, harsh events stand as the main index of driving aggressiveness. In order to detect aggressive driving tasks such as harsh acceleration, deceleration and cornering, the Oseven platform has developed sophisticated machine learning algorithms and data fusion techniques. Using accelerometer, gyroscope, and GPS related values, as input parameters, OSeven algorithms can detect harsh driving maneuvers with enough precision. It should be noted that the determination process of whether a driving maneuver constitutes a harsh event or not, does not rely on any predefined threshold or arbitrary rule, but instead has come as a result of a series of analyzes based on data fusion and advanced data mining techniques. The reliability of the identified harsh events has been repeatedly evaluated over the years both on scientific and commercial contexts.

4.1.2 The data sample

For the purpose of the specific research, the naturalistic driving database included 153,953 trips made from 696 unique drivers from December 2017 to August 2019. It should be highlighted that the dataset was provided in an anonymized format, since the only available information for each user was an identifier that it cannot be connected with any personal data. The trips were performed all around Greece, nevertheless the majority of them were conducted within the Region of Attica. For each trip, a variety of variables are available which include statistical measurements of acceleration and deceleration during a trip, speeding measurements that describe smoothly and with speed excess driving, as well as mobile usage indicators that describe how cautious the driver is. Table 4-1 presents the driving parameters used in the specific research.

Table 4-1. Driving parameters per trip

Variable	Description	Unit
harsh_acc_per_min	Average number of harsh accelerations performed per minute	events/min
acc_avg	Average acceleration	m/s^2
acc_std	Standard deviation of acceleration	m/s^2
acc_q90	90% percentile of acceleration	m/s^2
acc_max	Maximum acceleration	m/s^2
harsh_brk_per_min	Average number of harsh decelerations performed per minute	events/min
dec_avg	Average deceleration	m/s^2
dec_std	Standard deviation of deceleration	m/s^2
dec_q90	90% percentile of deceleration	m/s^2
dec_max	Maximum deceleration	m/s^2
speed_max	Maximum speed	km/h
mbu	Percentage of driving with mobile usage	%
speeding_percentage	Percentage of driving with speed over the speed limit	%

The main characteristics of the dataset as they emerged from the analysis performed herein, are given in Table 4-2.

Table 4-2. Main characteristics of the sample used in this research

	Total	Safe	Unsafe
Number of trips	153,953	66,566	87,387
Number of drivers	696	197	499
Average number of trips per driver		221	
Minimum number of trips per driver		16	
Average km travelled per driver		2,510 km	

Previous research has shown that there is a minimum amount of driving data that should be collected for each driver in order to obtain a clear picture regarding their own driving behavior (Stavrakaki *et al.*, 2020). This amount of necessary data depends on a number of parameters that include road type, driving aggressiveness, driving behavior volatility etc. An order of magnitude for the required number of kilometers traveled by a driver explicitly in an urban environment, in order to be able to study their driving behavior, is approximately 500 kilometers (Tselentis, 2018). The statistical characteristics of the total amount of distance traveled for the drivers in the sample are presented in Figure 4-5. The average total driving distance per driver in the sample is 2,510 km, corresponding to an average of 63 hours of driving, while the corresponding distance in urban environment is 917 kilometers.

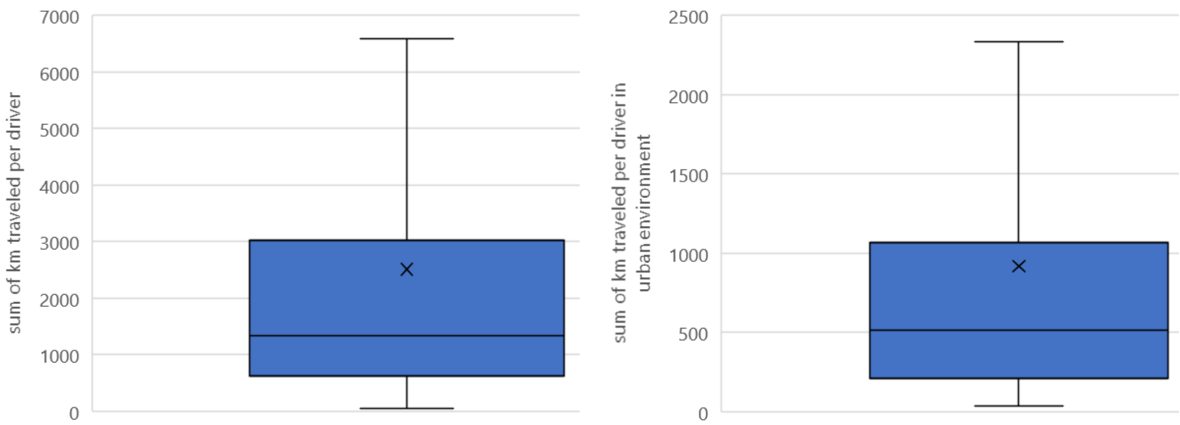


Figure 4-5. Boxplots of the distance travelled in kilometres in total (left) and explicitly in urban environment (right) per driver in the dataset

4.1.3 Sample representativeness

A significant innovation aspect of this dissertation is the development of a novel data-driven framework which aims to identify individual driving behavior and build a personalized driving decision support system. On the one hand, the data-driven approach allows to recognize real patterns straight from the data contrary to other approaches that use arbitrary assumptions. On the other hand, the results emerged from a data-driven analysis are inextricably linked with the specific sample, and only under specific conditions they can be generalized for the entire population. In most cases it is unclear whether the sample used to train the model is representative of the corresponding population and also whether its characteristics (driving behavior in the specific case) are similar to those of other samples of different population, e.g., drivers from other countries. Therefore, two main issues may arise when following a data-driven approach which refer to the generalization and transferability of the developed model and their outputs; yet the proposed methodological approach can still be transferable and easily interpretable to other samples as well, probably leading to different results.

When relevant data become available, indicatively including drivers' personal information such as gender, age and driving experience, it is vital to evaluate and tune the proposed recommendation system in order to ensure that unbiased outcomes are produced. Nevertheless, the dataset used in this work is user-agnostic and fully anonymized, and thus the representativeness of the sample could not be assessed.

4.2 Driving Profiles Identification

The first objective of this dissertation is to establish a methodological framework which can easily extract driving behavior profiles from raw smartphone sensed data. For this purpose, a clustering approach is implemented, and critical driving patterns are identified. All the necessary data processing and data filtering procedures as well as the clustering algorithm is implemented in R programming language using the RStudio environment.

4.2.1 Driving behavior per trip

The identification of driving behavior during a trip is based on a two-level clustering approach. Specifically, as also proposed in (Mantouka *et al.*, 2019), a two-stage k-means clustering is implemented on a selection of driving features such as: harsh events (acceleration and braking), acceleration and deceleration statistics (average, maximum and standard deviation), driving with excessive speed and distraction (mobile usage). The first level of clustering aims at isolating aggressiveness from the rest of unsafe driving patterns, as previous research has shown that although aggression is related with various unsafe driving habits, it does not necessarily imply an unsafe driving each self (Kockelman and Ma, 2018; Zahid *et al.*, 2020).

For the first level of clustering, the number of clusters is set to k=2 and clustering is implemented on Euclidean distance matrix. Two of the variables that are used for the above procedure describe the number of harsh alterations of the longitudinal position of the vehicle (acceleration and deceleration), while the rest of them are essentially indices of the average acceleration and deceleration of the trip. The results of this first implementation of the k-means clustering are presented in Table 4-3.

Table 4-3. 1st level clustering results

	Harsh acceleration per min	**Harsh brake per min**	**Average acceleration**	**Standard deviation of acceleration**	**Maximum acceleration**	**Average deceleration**	**Standard deviation of deceleration**	**Maximum deceleration**	**Number of trips**
Aggressive trips	0.150	0.2081	1.748	1.525	3.847	-1.968	1.843	-4.547	71263
Non-aggressive trips	0.028	0.051	1.137	1.052	2.503	-1.282	1.286	-2.926	82690

Based on the clusters' centers, the trips can be distinguished between aggressive and non–aggressive driving, since trips belonging to the first cluster are featured by aggressive driving characteristics, such as great acceleration and deceleration metrics and significantly higher rates of harsh events per minute of driving. The figure below (Figure 4-6) shows the relations between the variables used for clustering the data into two groups, as well as the way the observations are distributed between the two clusters.

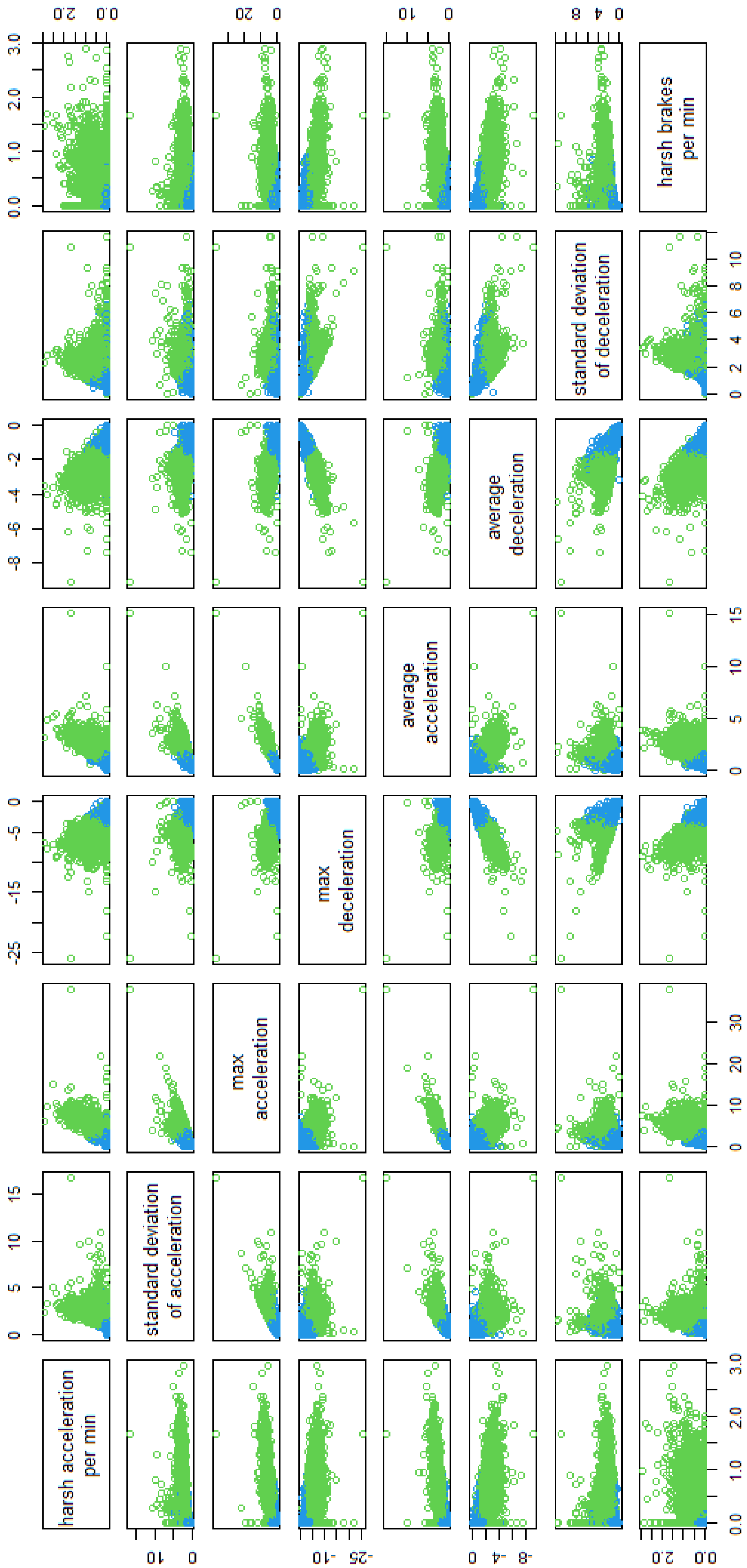


Figure 4-6. First stage clustering results: Clusters based on pairs of all variables

The performance of the clustering algorithm was assessed using the silhouette index which in this case was estimated 0.46 while the dunn's index was $4.18 \cdot 10^{-5}$, indicating that well-separated clusters were achieved (Figure 4-7).

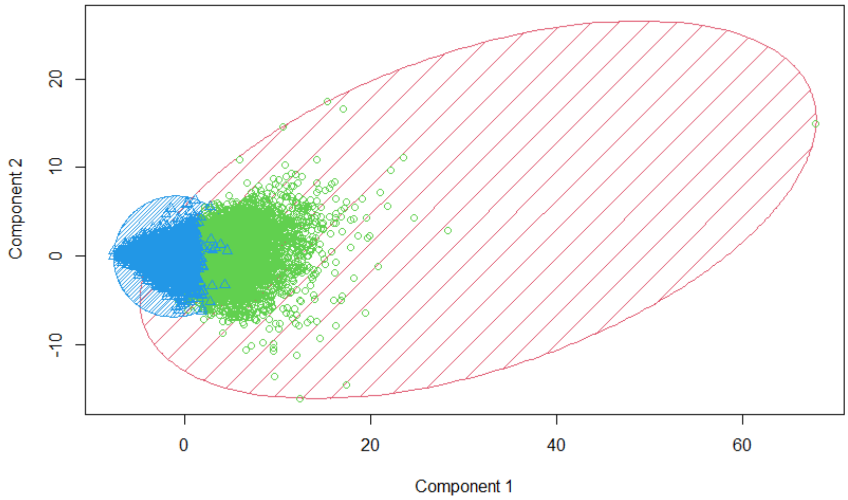


Figure 4-7. Depiction of first stage clustering results. The two components explain 82.37% of the point variability.

The second level of k-means clustering was applied separately to the two groups that emerged from the first level of clustering using two driving parameters: the percentage of driving with mobile usage and the percentage of driving with speed over the speed limit. Results of this second level of clustering are presented in the table below.

Table 4-4. 2nd level clustering results

	Percentage of mobile usage	Percentage of driving with speed over the speed limit	Number of trips
		Aggressive trips	
Distracted	0.511	0.062	4505 (2.9%)
Aggressive	0.019	0.032	54394 (35.3%)
Risky	0.023	0.269	12364 (8%)
		Non-aggressive trips	
Risky	0.021	0.306	12494 (8.1%)
Moderate	0.014	0.029	66566 (43.2%)
Distracted	0.514	0.057	3630(2.4%)

The resulting clusters seem to reveal richer driving profiles: distracted driving is recognized by higher values of the percentage of mobile usage while driving, while risky driving is identified through higher values of percentage of driving with speed over the speed limit. The two remaining clusters which have the lower values in both measures are annotated as "aggressive" and "moderate" for the aggressive and non-aggressive trips subsets respectively.

Regarding the performance of the second level of clustering, the silhouette index was estimated 0.52 in the case of the data subset that includes the aggressive trips and 0.55 for the non-aggressive trips, while the dunn's index was $5.57 \cdot 10^{-5}$ and $2.44 \cdot 10^{-5}$ respectively.

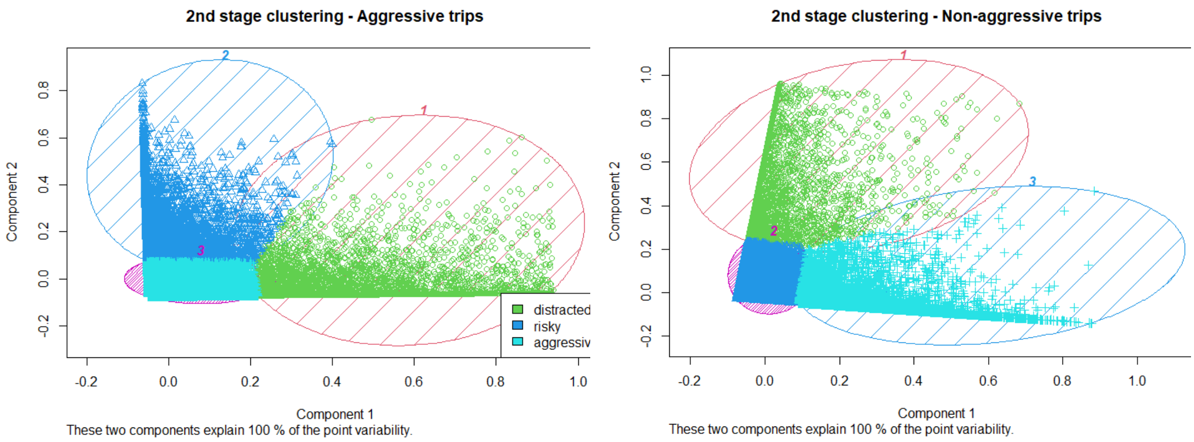


Figure 4-8. Depiction of second stage clustering results

4.2.2 Some discussion on driving behavior

The analysis performed on the real driving patterns highlighted the complexity of driving behavior and additionally, the variety of parameters that can define it. In this section, the most significant outcomes regarding the characteristics that govern driving behavior together with the differences that appeared between the identified driving profiles are discussed.

Findings revealed that, as shown in Figure 4-9, and in line with previous research, speeding behavior is much more severe in the case of non-aggressive trips, which indicates that aggressiveness does not necessarily imply an unsafe behavior such as driving over the speed limit (Kockelman and Ma, 2018). For this reason, some studies have distinguished between aggressive driving and driving that may be dangerous but not necessarily aggressive with regards to the driver's intentions (Richer and Bergeron, 2012; Zahid *et al.*, 2020).

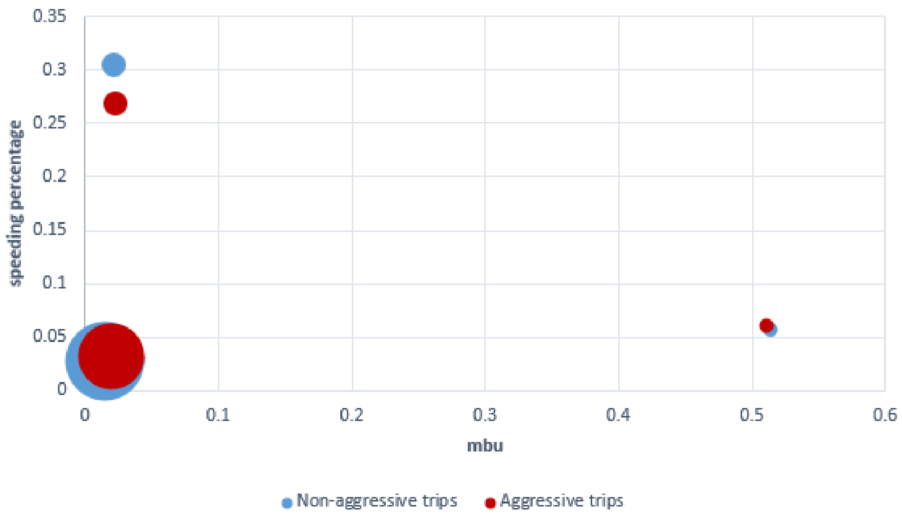


Figure 4-9. Cluster centers of the second level of clustering for both aggressive and non-aggressive trips

In order to give the complete picture of risk-taking while driving, the statistical characteristics of speeding percentage were calculated and they are presented for all trip profiles in Figure

4-10. Noteworthy differences can be observed between the profiles, where the greater values and amplitude of speeding percentage appears on the so-called "risky" profiles. Interestingly, risk taking is slightly impaired in the case where the drivers also exhibit aggressiveness (aggressive-risky profile) as well.

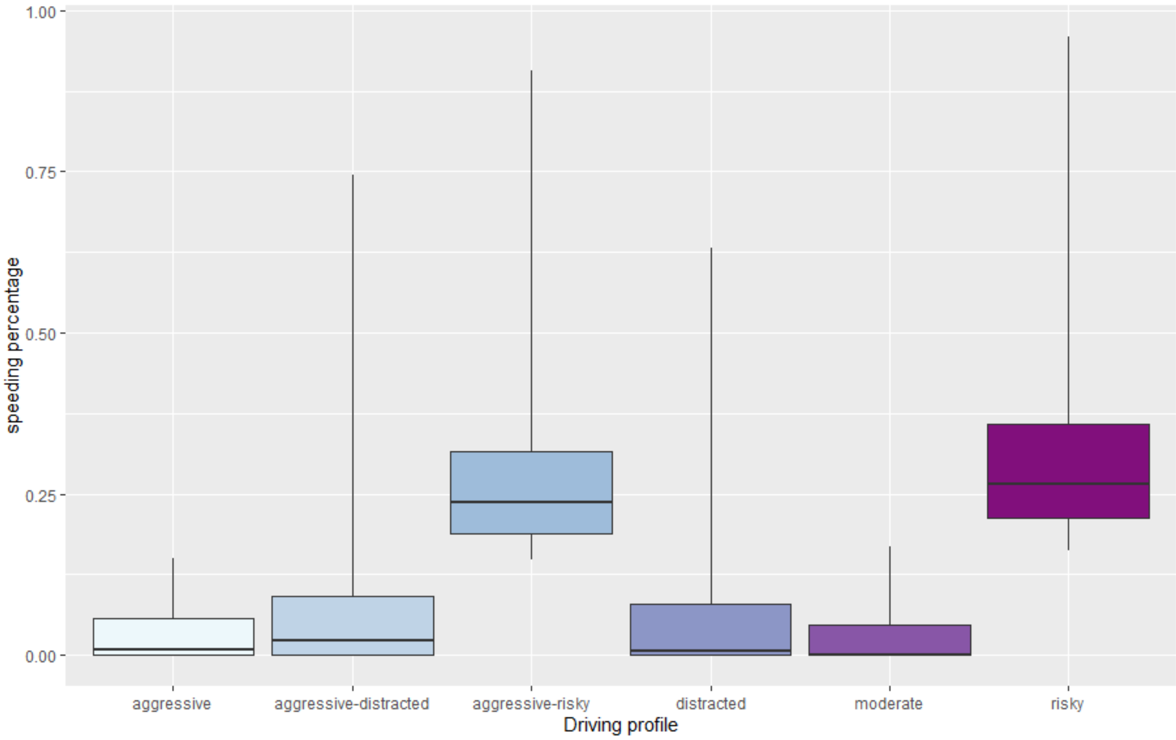


Figure 4-10. Boxplot of speeding percentage for each driving profile

Figure 4-11 shows the descriptive statistics of the average acceleration of the six identified profiles. Significant differences appear especially when comparing aggressive with non-aggressive driving profiles. The three aggressive profiles have greater average acceleration metrics, while non-aggressive profiles present a more condensed boxplot. Such findings indicate that acceleration decisions can be considered as the most appropriate indicator of one's personal driving style as they are only dependent on the driver's perception and preference between smoothly or harshly accelerating. Indeed, in recent literature, a driver's driving style is usually defined by their acceleration profile (Zhang *et al.*, 2020; Fafoutellis *et al.*, 2021). Liu *et al.* (2019) claimed that studies on acceleration behavior could enhance human-like driving ability of the driving automation systems and autonomous vehicles.

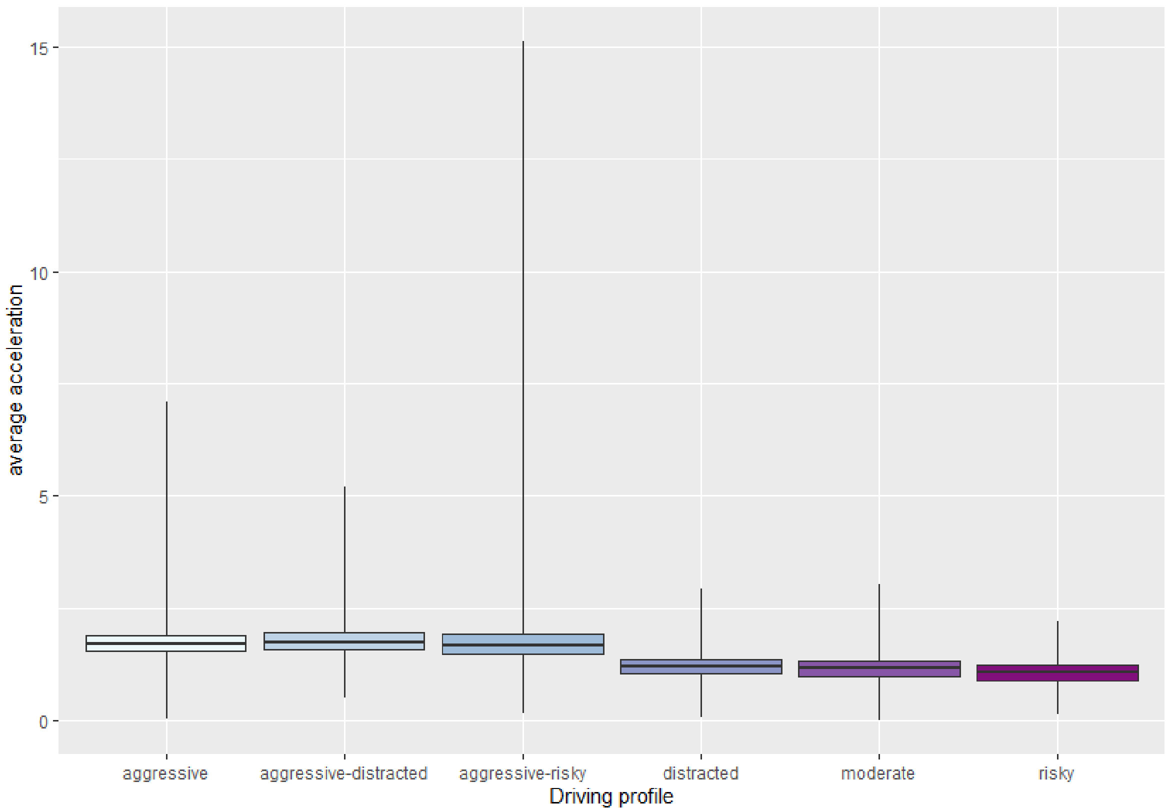


Figure 4-11. Boxplot of average acceleration for each driving profile

On the contrary, the selection of the appropriate speed is highly correlated with the road geometry and road traffic. More specifically, the road speed limit value has a significant impact on driving speed, and thus, a driver will choose to increase or decrease speed based on this value (Hamzeie *et al.*, 2017; Liu *et al.*, 2020). Previous research has also shown that other critical determinants of drivers' speeding behavior are the road illumination, the presence of horizontal curves and longitudinal slope changes (Sadia *et al.*, 2018). Other researchers also have shown that speed choice is affected by route familiarity and more precisely, speed increases with the repetition of driving on the same route (Colonna *et al.*, 2016).

Finally, deceleration decisions are usually dependent on the leading vehicle's behavior and traffic signals. A great body of literature has analyzed drivers' deceleration behavior at signalized intersections, with a great focus on the onset of a yellow-phase transition (El-Shawarby *et al.*, 2007, 2011; Rittger *et al.*, 2015). More recent literature has also developed emergency braking systems for electric and autonomous cars, which are able to detect the environment and perform emergency decelerations if necessary (Cicchino, 2017; Min *et al.*, 2019). This fact implies that in some cases deceleration decisions are highly correlated with exogenous parameters, such as the existence of other road users (pedestrians, cyclists, etc.), the behavior of the leading vehicle, unexpected events and many more (El-Shawarby *et al.*, 2007; Angkititrakul *et al.*, 2009).

4.2.3 Average driving behavior per driver

In order to separate drivers into groups with the same driving preferences, an average driving profile of each individual was identified by applying a simple rule. All four driving profiles indicating an unsafe driving behavior (Risky, Distracted, Aggressive-risky, Aggressive-distracted) were grouped as the worst class (3), aggressive trip profiles constitute the second class (2), while trips with typical characteristics belong to the first class (1), as shown in Figure 4-12. For each individual driver, an average from all their trips is estimated and drivers are separated into two main groups based on their average behavior, as follows:

- Moderate/typical drivers: *trip average* $\leq$ *1.5*
- Reckless drivers: *trip average* > *1.5*

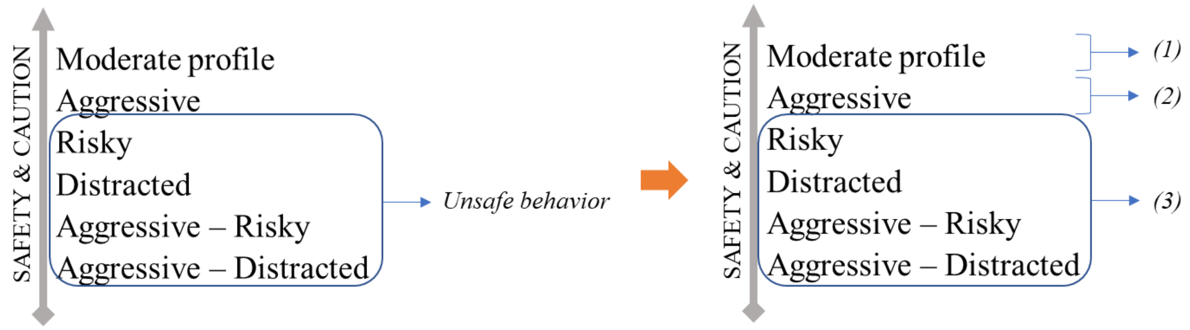


Figure 4-12. Trip profile grouping for drivers' average driving profile estimation

For each individual driver, an average of the annotations from all their trips is estimated, where trip average less than 1.5 implies a moderate/typical driver and trip average greater than 1.5 refers to reckless drivers. Based on the statistics presented in Figure 4-13, trip average less than 1.5 indicates that at least 60% of the trips performed by a driver are characterized by "moderate" driving behavior. In order for the developed controller to be as adaptive as possible to each individual's behavior, the proposed framework should be very strict when characterizing a driver as "typical/moderate" in order to avoid suggesting changes in behavior that the driver himself is impossible to follow as they will be far from his own average behavior.

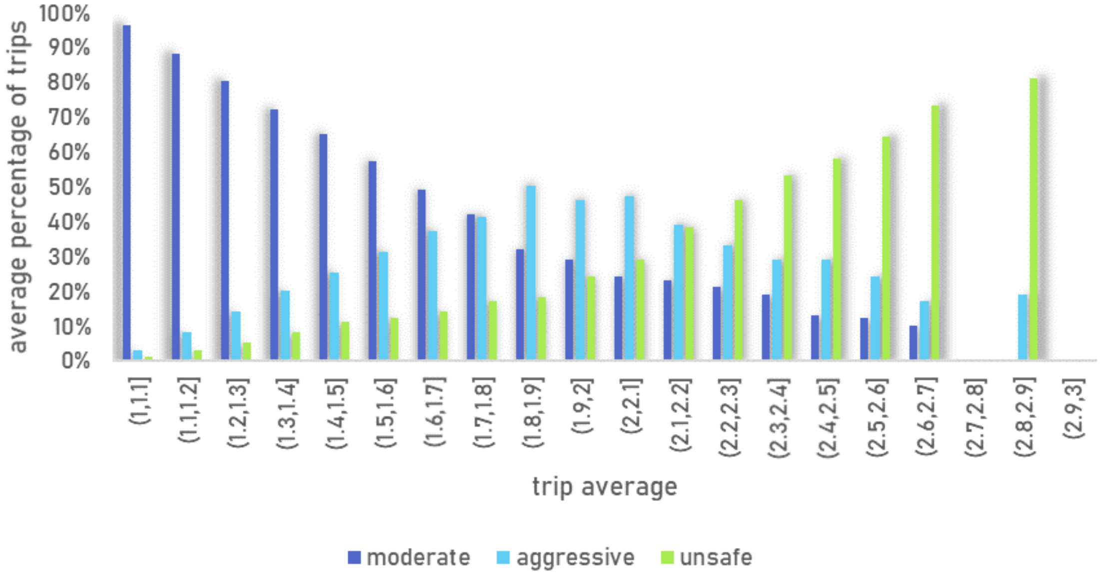


Figure 4-13. Average percentage of trips of each driving profile for different ranges of drivers' trip average score

4.3 Building the RL Controllers

The recommendation system proposed within this dissertation is basically a decision support system for drivers that aims at mitigating aggressiveness and riskiness. Its development is based on a Reinforcement Learning algorithm which is capable of producing the optimal behavior alteration for each driver given the way they have drove over their last trip.

4.3.1 Conceptual design

Driving is a complex task since it requires from the driver to take both strategic and dynamic decisions as well as adapt their behavior to emerging conditions of the network. Contrary to the already developed ADAS, the system here has the following three state-of-the-art characteristics:

1. It is *personalized*, which means that it recommends the best driving actions to each individual taking into account their specific requirements and driving preferences.
2. It is *self-aware*, which means that the system takes into account previous behavior of each individual driver in order to propose the most suitable driving recommendations.
3. It is *autonomous*, meaning that it does not require any external input from the network or the traffic. Driving recommendations aim to improve individual driving behavior on its core, namely acceleration and deceleration decisions.

In order to develop the Self-Aware Driving Recommendation Assistant (SADRA) which is proposed here, a structured procedure is followed. First, the total trip database is divided into two, based on the average driving profile of each driver. In particular, the first database includes the trips of all drivers belonging to the "typical-safe" drivers, while the second includes all the trips of drivers with unsafe average driving behavior. For the sake of brevity, from this point on, the RL controller that corresponds to the “typical” drivers is referred to as SADRA – I, while the corresponding controller for the reckless drivers is referred to as SADRA – II respectively.

4.3.2 Problem setup

Every RL agent consists of three main components: states, actions and rewards. In each timestep the agent observes the current state of the environment and takes the appropriate action from the set of the possible actions. Then, the agent receives a reward which measures the success or failure of the agent’s actions for the given state.

In this study, the environment states are defined through a five-dimensional vector that describes how a driver drove during their trip and includes trip’s average acceleration (a_{avg}), 90% percentile of acceleration (a_{90}), average deceleration (d_{avg}), 90% percentile of deceleration (d_{90}) and percentage of driving with speed over the speed limit (*speeding*):

$$s = \{a_{avg}, a_{90}, d_{avg}, d_{90}, speeding\} \quad (15)$$

Our recommendation system is not context-aware which means that its ultimate goal is to improve individual’s personal driving style independently from the road setting they are driving in (type of road, traffic conditions, etc.). The selection of the appropriate speed is not independent from the road geometry and road traffic, as well as deceleration decisions are not always independent from the leading vehicle’s behavior and traffic signals. Therefore, the only parameter that purely describes one’s driving style is the acceleration, as it is only dependent on the driver’s perception and preference between smoothly or harshly accelerating. Indeed, in recent literature, a driver’s driving style is usually defined by their acceleration profile (Zhang *et al.*, 2020; Fafoutellis *et al.*, 2021). To this end, actions that the system produces and are proposed to the driver belong to a continuous action space which is defined by a two-dimensional vector including a change in average acceleration and in the 90% percentile of acceleration, which define the usual/preferred acceleration for the entire trip in regular situations and the value that should not be exceeded, e.g., when performing overtaking maneuvers, except from cases of emergency:

$$a = \{da_{avg}, da_{90}\} \quad (16)$$

For the sake of simplicity from hereon, the 90% quartile of the acceleration may be equally referred to as “maximum acceleration”.

A key component of the RL agent is the reward function. The aim of the reward function is twofold; to evaluate the current state and the transition between states. In other words, the driving behavior at each trip, as well as the change in driving behavior between successive trips of the same user are evaluated. For this purpose, a custom driving evaluation function had to be constructed first. The score of each trip was estimated by the distance of this specific trip from the center of the moderate profile (the center of the cluster), in order to quantify how far each individual's behavior is from the typical (moderate) behavior. The values of each driving behavior parameter that correspond to the "moderate" profile are shown in Figure 4-14, together with the minimum and maximum values of each parameter as they were observed in the data.

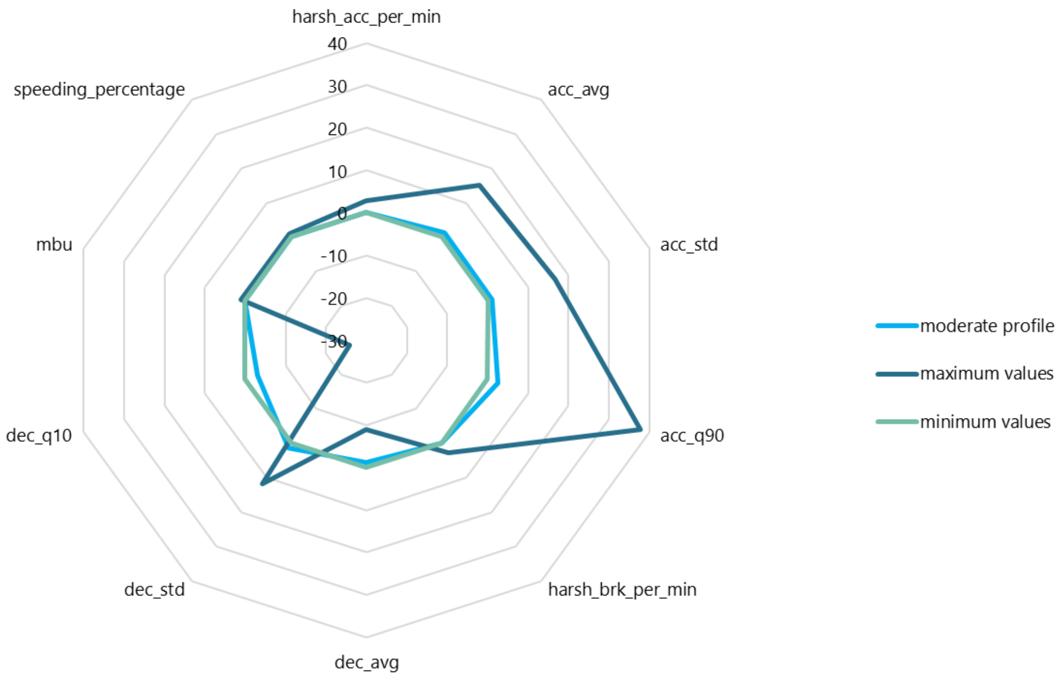


Figure 4-14. Typical driving profile center in relation to the minimum and maximum values of the driving parameters.

For the purpose of this analysis, the Mahalanobis distance is used to estimate the distance between each trip and the moderate profile. The Mahalanobis distance between two objects is defined as follows (Varmuza and Filzmoser, 2016):

$$\text{Ϻ} = \sqrt{[(x_B - x_A)^T * C^{-1}(x_B - x_A)]} \qquad (17)$$

where $x_A$and x_B is a pair of objects, and C is the sample covariance matrix. In contrast to the Euclidean distance, the Mahalanobis distance takes into account the correlation structure of the data as well as the individual scales (Barnett and Lewis, 1994).

The 75% of trips abstain up to 11.11 from this moderate profile, while there is a small number of trips that appear to have extreme distances from this average driving behavior. The statistics of the estimated Mahalanobis distance are presented in Table 4-5.

Table 4-5. Descriptive statistics of the mahalanobis distances

Statistic measure	Value
Minimum	0.13
1st quartile (25%)	2.96
Median	5.57
Mean	10.67
3rd quartile (75%)	11.11
Maximum	3677.81
skewness	44.76
kurtosis	3545.31
se	0.08

The estimated mahalanobis distances (Figure 4-15) are exploited in an exponential function, where the more negative the exponent, the steeper the graph and consequently the more unsafe behavior the more negative the trip score is. Consequently, using such a formula, the difference between the profiles is very well-defined.

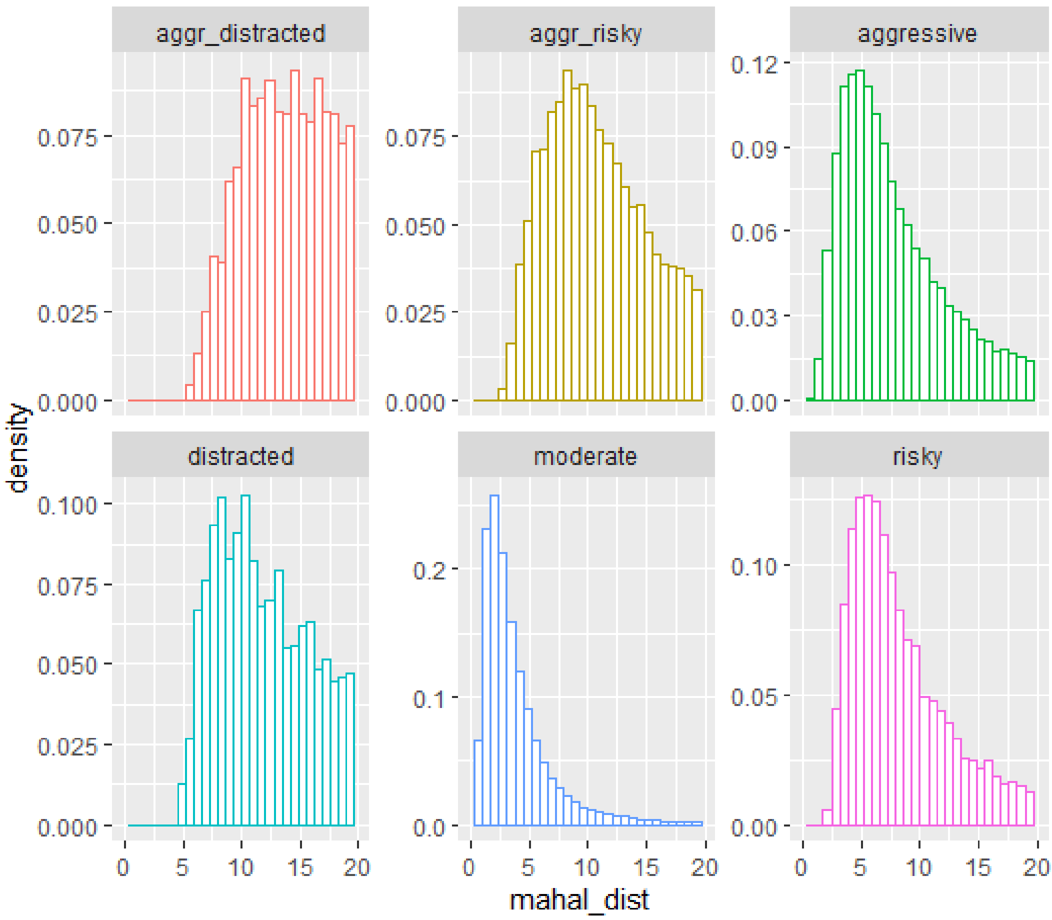


Figure 4-15. Mahalanobis distance distribution per trip profile

Finally, trip evaluation is performed on the basis of the following formula:

$$trip\ score_i = e^{-driving\ profile_i * \frac{M_{(i,moderate\ profile)}}{Q_{75}(M)}} \tag{18}$$

where i is an individual trip and Ɱ is the Mahalanobis distance. Here, the 3rd quartile of the Mahalanobis distance is used instead of the maximum value in order for the score function to be stricter with drivers whose behavior excludes more than 75% of the typical (moderate) behavior. A graphical representation of the score function is given in Figure 4-16.

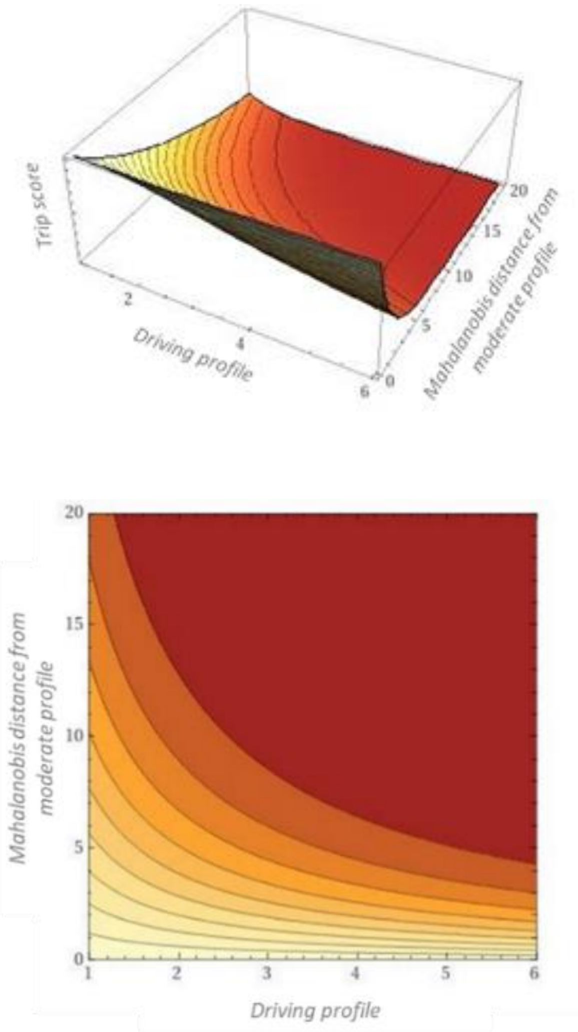


Figure 4-16. 3D and 2D graphical representation of the score function

The score is scaled from 1 – 100 so that it can be easily interpreted within the rest of the algorithms. The greater the trip score the better (less aggressive and risky) the driving behavior of the specific trip is. The distribution of the trip score per driving profile is illustrated in Figure 4-17.

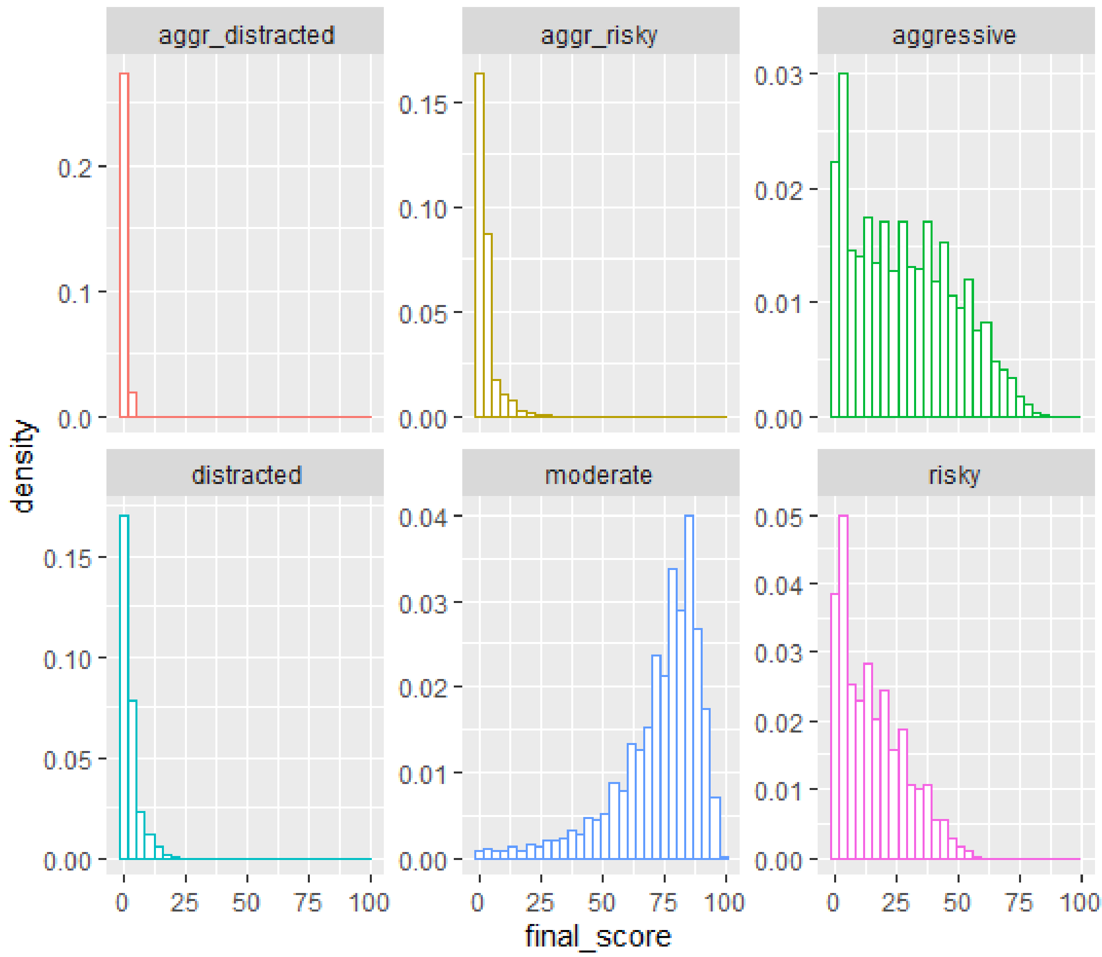


Figure 4-17. Total trip score distribution per trip profile

Finally, the reward function for a driver moving from one trip to the next one was established based on the following formula:

$$r = \ trip\ score_{i+1}\left(1 + \frac{trip\ score_{i+1} - trip\ score_i}{100}\right) \tag{19}$$

Once the main components for the development of the RL controllers were estimated, the data were organized in the following format:

$$(state, action, reward, next\ state) \tag{20}$$

For every unique driver in the dataset, their trips were sorted in an ascending order according to each trips starting date. The training samples were tuples of sequential trips of a specific driver along with the corresponding action and reward of the transition from the first trip to the succeeding one. It should be noted that for every distinct driver in the dataset, their first trip was used only as "*state*" while their last trip of was used only as "*next state*". Following this data preparation procedure, 33,440 unique data samples were constructed for training SADRA I and 119,817 unique data samples were used for the training process of SADRA II.

Data preparation and filtering are performed in Python programming language and several scripts are written in order to prepare the data in the appropriate format for training the RL agents. Reinforcement Learning algorithms are also developed in Python programming language. Coding is applied using Anaconda Environment (Spyder), for Python & scientific development. The computer used for the computation time estimation is an Intel® Core™ i7 CPU K 875 @ 2.93GHz × 8 featuring a 32.0 GB Ram memory running on Windows 10Pro.

4.3.3 Model development

In Deep Learning, searching for the optimal structure of a Neural Network model and tuning its hyperparameters is a very crucial task that may significantly improve the model's performance. This section presents the results obtained after performing an exhaustive grid search to conclude to the final architecture of the two networks, namely actor and critic. Specifically, all possible combinations of the networks' structures and parameterization, within a range of reasonable values, have been examined and compared in order to detect the optimal one. The parameters that were taken into consideration are: number of hidden layers, number of neurons and activation of each layer, optimization algorithm and learning rate, batch size and number of training epochs.

4.3.3.1 Typical drivers' controller – SADRA I

The RL controllers are developed based on the DDPG algorithm which implements an actor-critic approach to learn a policy and produce the optimal actions. Thus, for each controller two neural networks are developed; representing the actor and the critic respectively. In this section, the structure of the first version of the RL controller (SADRA I) that corresponds to drivers with a typical/safe average behavior is given.

First, instead of starting with a random network, an initial training of the networks is performed using only the rewards. Then, both the critic and actor networks are trained using the hyperparameters presented in Table 4-6, emerged after an exhaustive grid search. The results of the grid search over the hyperparameters of the neural networks are presented in the figures below, both for the critic (Figure 4-18, Figure 4-19) and the actor network (Figure 4-20, Figure 4-21).

Critic network

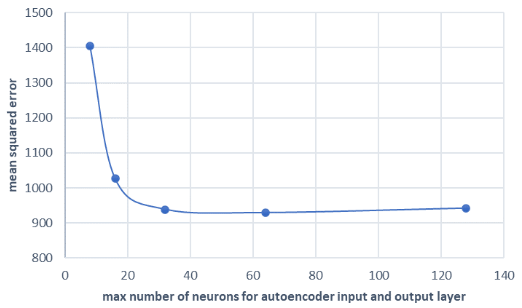


Figure 4-18. Validation loss versus max number of neurons for critic network – SADRA I

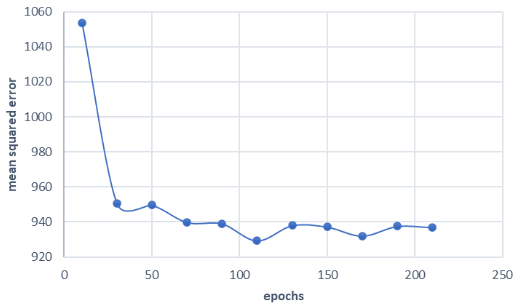


Figure 4-19. Validation loss versus number of epochs for critic network – SADRA I

Actor network

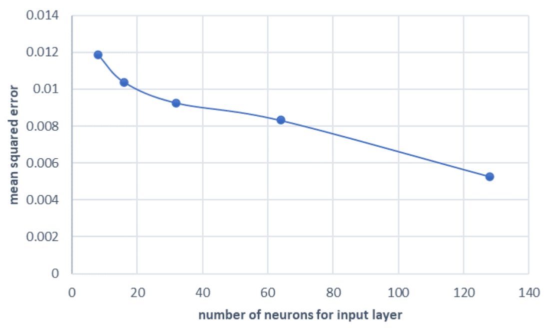


Figure 4-20. Validation loss versus max number of neurons for actor network – SADRA I

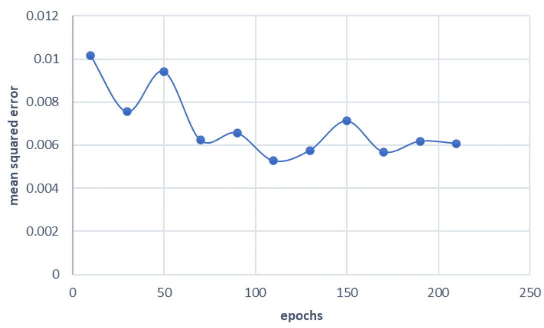


Figure 4-21. Validation loss versus number of epochs for actor network – SADRA I

The critic network, which estimates the Q-value for any given pair of (state, action), has an autoencoder-like architecture with six hidden layers, a 7-units input layer and 1-unit output layer, as depicted in Figure 4-22. The actor network, which estimates the best possible action for any given state, consists of three hidden layers, a 5-units input layer and a 2-units output layer. The rectifier linear unit activation function (ReLU) has been used for the neurons of all layers, for both networks. The model was fitted using the Adam optimizer with a learning rate equal to 0.001 for the critic network and 0.0001 for the actor network, for a training period of 200 epochs and a batch size of 150.

Table 4-6. Hyperparameters of Critic and Actor Networks (SADRA-I)

Hyperparameters	Critic network	Actor network
Number of hidden layers	6	3
Number of neurons per layer	(64,32,16,16,32,64,1)	(128,64,32,2)
Epochs	200(initial network:110)	200(initial network:110)
Batch size	150(initial network:150)	150(initial network:150)
Activation	ReLU	ReLU
Optimizer	Adam	Adam
Learning rate	0.001	0.0001

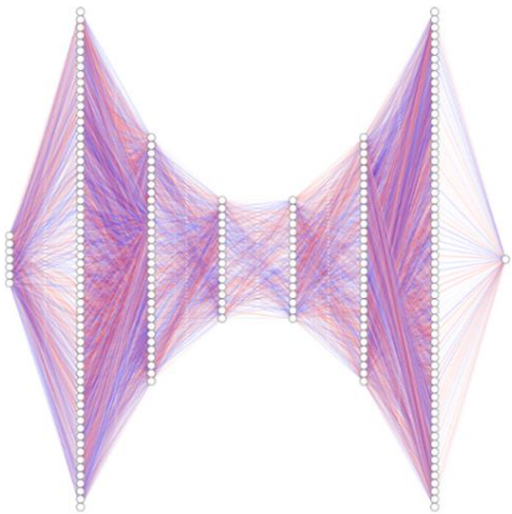


Figure 4-22. Architecture of the Critic Network – SADRA I

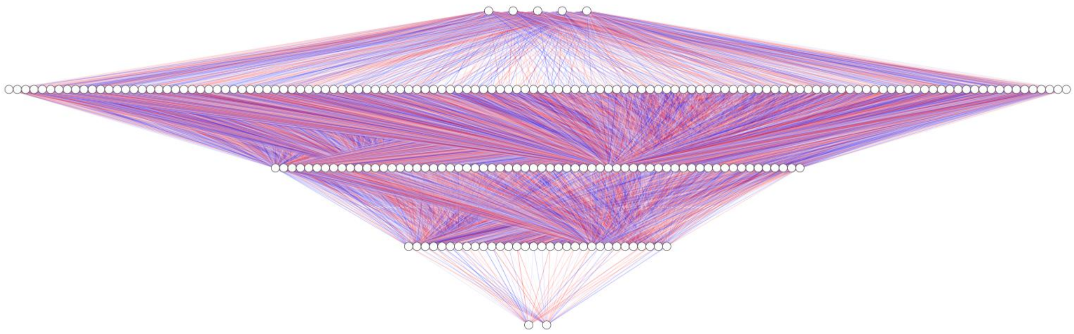


Figure 4-23. Architecture of the Actor Network – SADRA I

Once the hyperparameters for both the critic and the actor networks are defined and the structure of the two neural networks is formed, the DDPG model is trained using 33,440 data samples corresponding to arrays of successive trips performed by 197 unique drivers.

The performance of the neural networks is evaluated using the Mean Squared Error as an evaluation metric. Results shown in Figure 4-24 reveal that the two networks converge after a number of epochs.

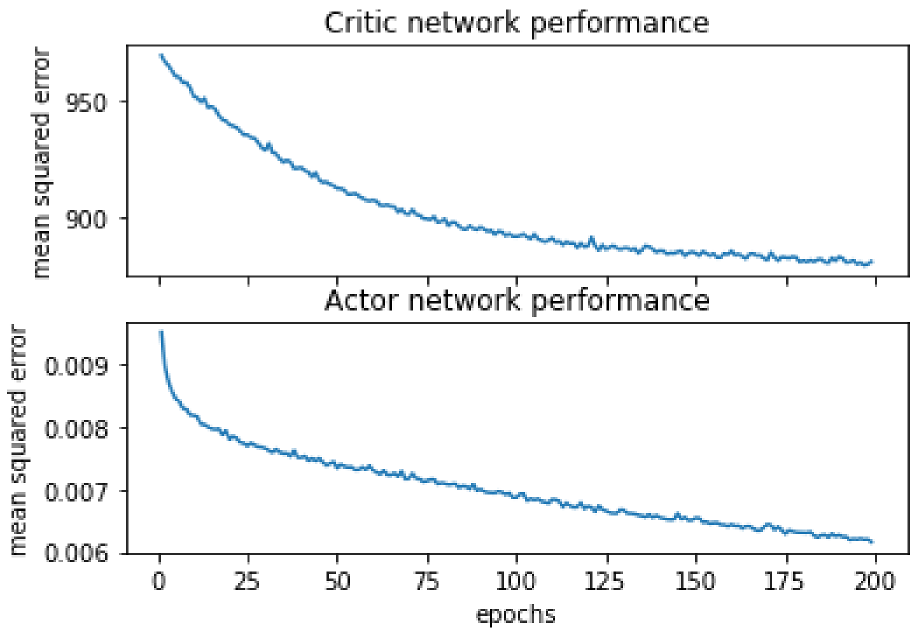


Figure 4-24. Critic and Actor network performance corresponding to SADRA I

The trained model of this first version of the RL controller (SADRA I) can be used to generate driving recommendation for drivers that exhibit a safe overall driving behavior. The results of the controller are discussed later on, in comparison with the outcomes of the unsafe drivers' controller.

4.3.3.2 Unsafe drivers' controller – SADRA II

In this section, the structure of the second RL controller is given, which corresponds to drivers that perform a variety of unsafe driving behaviors during their trips.

Figure 4-25 and Figure 4-26 depict the loss curves of the critic network for SADRA II, that corresponds to the unsafe drivers' RL controller, as they emerged after the grid search performed over the hyperparameters of the network. Based on the results presented in Figure 4-25, the maximum number of neurons used on the first and last layer of the autoencoder structure of the critic network is 32 and as depicted in Figure 4-26, the appropriate number of epochs is 170. The rectifier linear unit activation function (ReLU) has been used for the neurons of all layers, for both networks. The model was fitted using the Adam optimizer with a learning rate equal to 0.0001. It should be noted that the hyperparameter tuning was performed for the initial network which was trained based solely on the rewards.

Critic network

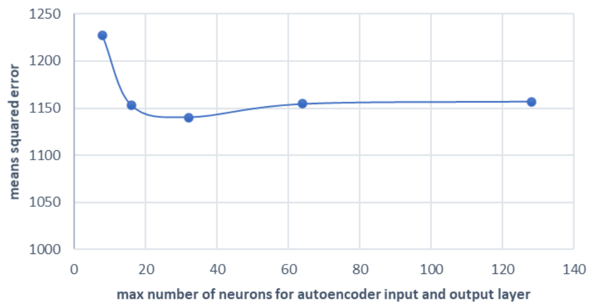


Figure 4-25. Validation loss versus max number of neurons for critic network – SADRA II

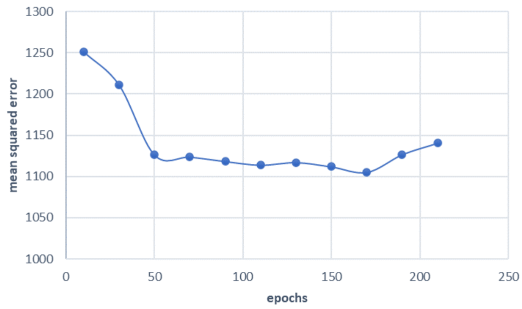


Figure 4-26. Validation loss versus number of epochs for critic network – SADRA II

The corresponding loss curves for the actor network are depicted in Figure 4-27 and Figure 4-28 for the maximum number of neurons in the input layer and the number of epochs respectively. The final architecture of the actor network is shown in Figure 4-30.

Actor network

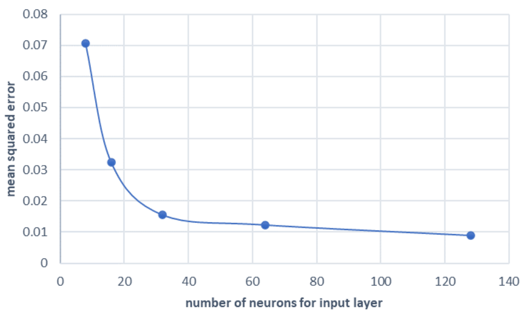


Figure 4-27. Validation loss versus max number of neurons for actor network – SADRA II

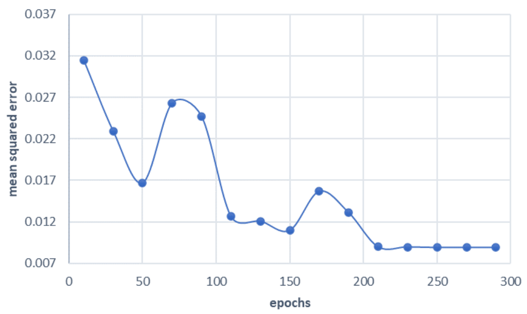


Figure 4-28. Validation loss versus number of epochs for actor network – SADRA II

According to the hyperparameter tuning conducted through the grid search, the appropriate hyperparameters for the actor network are as follows: maximum number of neurons for the input layer 128 and number of epochs 210. The neural network was fitted using the Adam optimizer with a learning rate equal to 0.0001 and a batch size of 250. The corresponding hyperparameters of SADRA II are summarized in Table 4-7.

Table 4-7. Hyperparameters of Critic and Actor Networks (SADRA II)

Hyperparameters	Critic network	Actor network
Number of hidden layers	6	3
Number of neurons per layer	(32,16,8,8,16,32,1)	(128,64,32,2)
Epochs	170(initial network:170)	210(initial network:210)
Batch size	250(initial network:250)	250(initial network:100)
Activation	ReLU	ReLU
Optimizer	Adam	Adam
Learning rate	0.0001	0.0001

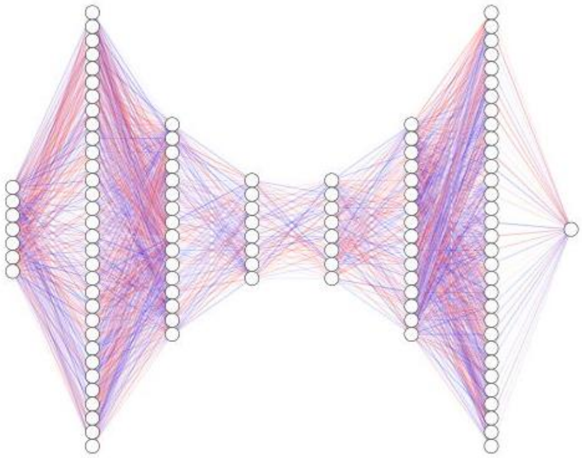


Figure 4-29. Architecture of the Critic Network – SADRA II

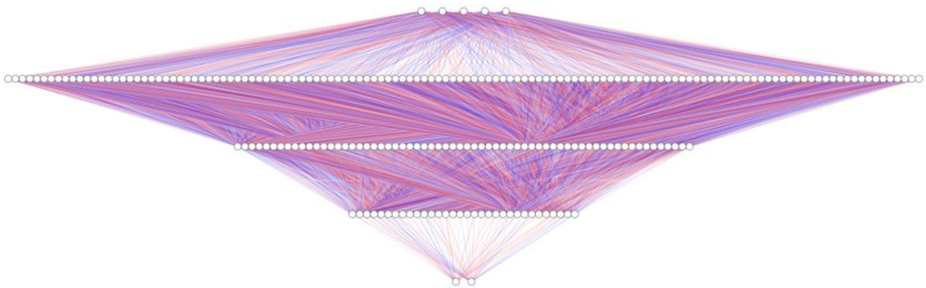


Figure 4-30. Architecture of the Actor Network – SADRA II

SADRA II was trained using 119,817 data samples corresponding to arrays of successive trips performed by 499 unique drivers. Using as loss function the Mean Squared Error it can be observed that both networks converge after a number of epochs (Figure 4-31).

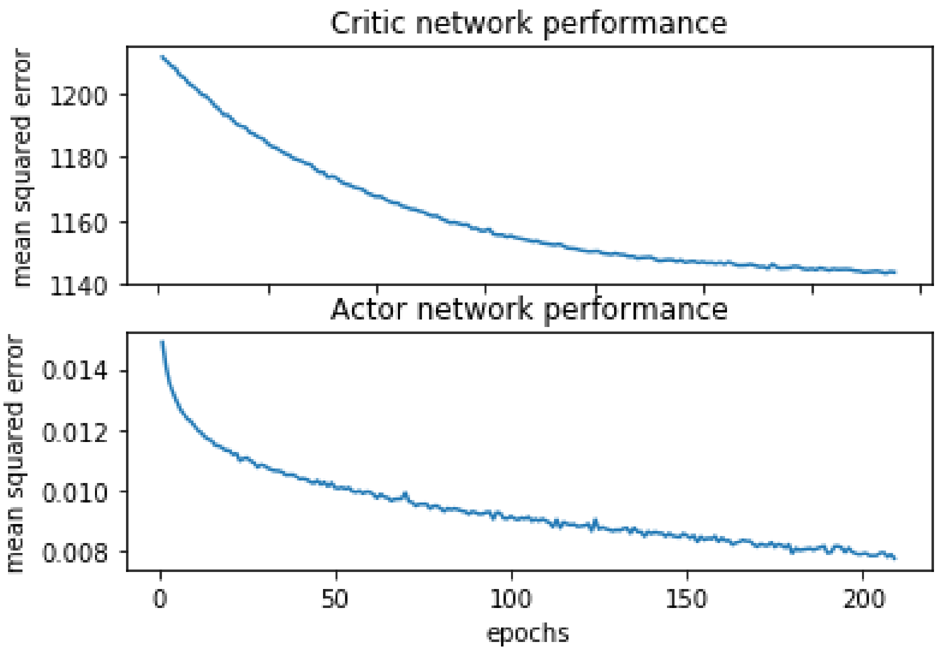


Figure 4-31. Critic and Actor network performance corresponding to SADRA II

The actor and critic networks for both the safe and unsafe drivers' subsets were trained following the procedure of Algorithm 1.

Algorithm 4: DDPG Algorithm implementation

Initialize critic $Q(s,a|\theta^Q)$ and actor $\mu(s|\theta^\mu)$ networks using rewards as Q-values
Set the above as initial target networks (Q' and μ')
Split the sample into M minibatches
for minibatch=1, M **do**
  Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$
  Update critic by minimizing the loss: $L = \frac{1}{N}\sum_i \left(y_i - Q(s_i, a_i|\theta^Q)\right)^2$
  Update the actor policy using the sampled policy gradient:

$$\nabla_{\theta_\mu} J \approx \frac{1}{N}\sum_t [\nabla_a Q(s,a|\theta^Q)|_{s=s_t, a=\mu(s_t)} \nabla_{\theta_\mu}\mu(s|\theta^\mu)|_{s=s_t}]$$

  Update the target networks:

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$$
$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$$

end for

4.4 Simulation setting

The quantification of the impact of adopting driving recommendations by all drivers on traffic, road safety and emissions was performed under a network-level microscopic simulation scenario. Simulation is a powerful tool for road traffic analysis and prediction. Microsimulation refers to the tracking of individual vehicle's movements on a second or sub-second basis. Most of the traffic simulation tools are based on the idea that the speed of a vehicle depends on the speed of the leading vehicle, which is the main idea of car-following models (Bieker-Walz *et al.*, 2017).

Here, the SUMO simulation software is used and its default car-following model, Krauss model. SUMO is an open-source microscopic road traffic simulation tool, which consists of several supportive applications (e.g., netconvert and netgenerate) designed to prepare the simulation scenario (Krajzewicz et al., 2012). It implements the car-following model of Krauss, a microscopic, space-continuous model based on the safe speed; the driver of the following car adopts a safe speed which allows them to adapt to the deceleration of the leading vehicle (Krajzewicz et al., 2005).

4.4.1 The case study of Athens

The case study for the simulation experiments is the inner-ring network of Athens, Greece (Figure 4-32). The network was exported from the Aimsun simulation software and then imported on the SUMO microscopic simulator, with all the necessary changes. The network consists of 1,293 nodes/intersections and 2,572 edges. The total length of the network is 348 kilometers.

Figure 4-32. The Athens' Road network

Concerning the edges, 13 different edge types have been used, in order for the speed limits and the priority of the edges to be imported in SUMO. Regarding traffic lights, 440 traffic lights have been inserted manually on the Athens inner-ring network. The information about the traffic lights concerns the traffic light programs corresponding to the morning peak hour. In order for the network to be calibrated correctly, the following have been extracted from the Aimsun software:

- Frank-Wolfe User Equilibrium path assignment results

- Counts for all the edges of the network in the peak hour
- Total number of vehicles in the network during peak hour (approx. 86,000 vehicles)

Said elements from the Aimsun simulator led the trip distribution procedure on the SUMO microscopic simulator. The calibration of the network led to the definition of **86,054 vehicles**, achieving a total of 1,393,634 counts (97.47% of the total counts extracted from the Aimsun simulator) and a GEH value below 5 (GEH < 5) for 95.26%.

4.4.2 Scenarios setting

Two distinct scenarios were designed both corresponding to the demand of the Athens' Road network during the morning peak hour (8:00 – 9:00 AM). First, the initial conditions of the network are simulated in order to estimate the performance of traffic when vehicles move around, based on the characteristics that govern the six identified driving profiles. In order to ensure the robustness of the results, simulation was performed in 10 replications with ten different seed numbers. Stochasticity is an important aspect of reproducing reality in a simulation scenario, since it adds randomness over the distributions of difference aspects of the simulation (e.g., route distributions, vehicle type distributions).

Subsequently, driving recommendations were produced offline for every served vehicle based on the way each vehicle performed their trip. The recommendations were produced from the corresponding RL controllers using as input the state of the trip (average acceleration, 90% percentile of acceleration, average deceleration, 90% percentile of deceleration, speeding percentage) and as output the optimal alteration of the maximum acceleration. It should be highlighted here that although the developed RL controllers produce a two-dimensional vector that includes alterations on both the average and the maximum acceleration, only the maximum acceleration was exploited during the simulation runs, since the Krauss model takes into account only the maximum values of acceleration and deceleration.

Finally, a second simulation run was performed, where previously served vehicles follow the proposed recommendations, namely an alternation of their maximum acceleration, while the rest of the traffic follows the distribution among the six driving profiles.

The behavior that implies each driving profile was simulated through the adjustment of the car-following model. The car-following model can be parametrized by a number of parameters: the maximum acceleration of the vehicle (*accel*), the maximum deceleration of the vehicle (*decel*), the maximum velocity of the vehicle (*maxSpeed*), the maximal physically possible deceleration for the vehicle (*emergencyDecel*) and the vehicles' expected multiplicator for lane speed limits (*speedFactor*). At first, the current (initial) state of the road traffic is simulated in SUMO using the six defined driving profiles, whose parameters were introduced to the Krauss model of different vehicle types, as shown in Table 4-8.

Table 4-8. Car-following model parameters for each vehicle type

Vehicle types (trip profiles)	Car-Following Model Parameters accel (m/s^2)	decel (m/s^2)	emergencyDecel (m/s^2)	maxSpeed (km/h)	speedFactor (mean, min, max)
Moderate	2.519	-2.942	-5.909	64.51	(0.029, 0, 0.168)
Aggressive	3.817	-4.483	-18.083	66.93	(0.033, 0, 0.151)
Risky	2.392	-2.824	-5.328	100.28	(0.306, 0.1627, 0.96)
Distracted	2.601	-2.990	-5.112	67.38	(0.057, 0, 0.631)
Aggressive-risky	3.944	-4.825	-25.884	100.8	(0.269, 0.147, 0.907)
Aggressive-distracted	3.939	-4.553	-10.845	71.99	(0.062, 0, 0.744)

The speed factor follows a distribution based on the mean, min and max values, as they emerged from the clustering analysis. The length of a vehicle is set to 4.5 m and the minimum net gap between the leader and the follower is set to 2.5 m. Each vehicle type was also depicted with a different color, as shown in Figure 4-33.

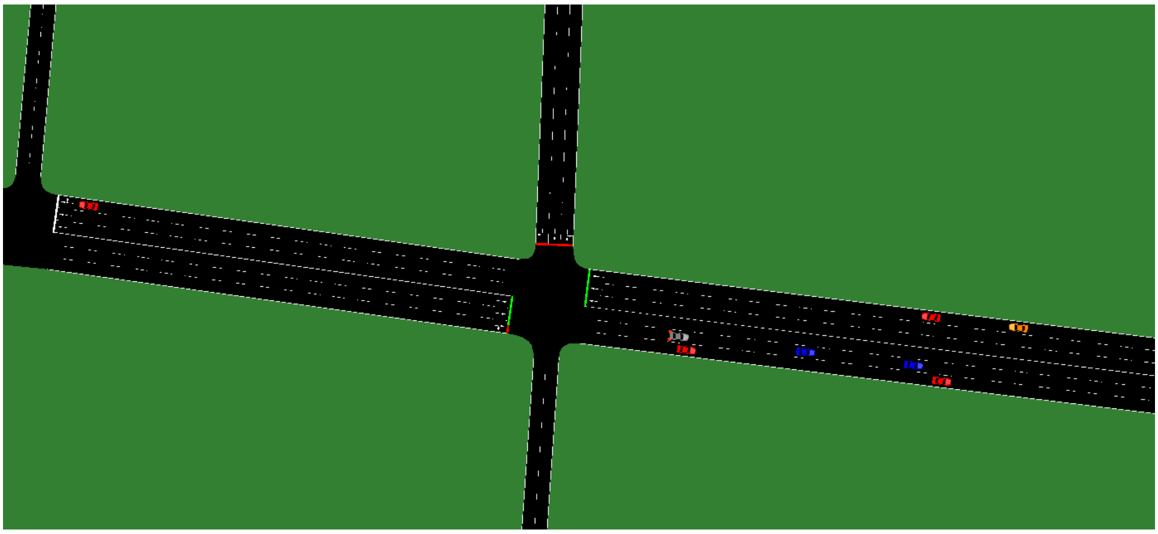


Figure 4-33. A screen capture of the simulation software depicting the different type of vehicles with different colors

One of the main components of SUMO simulation tool is the route file, which holds a variety of information about the characteristics of the demand and specifically, the vehicle type which describes the vehicle's physical properties, the route each vehicle shall take, and the vehicle itself which gives information, among others, on the vehicle type, the departure time and the id of each vehicle. Both routes and vehicle types can be shared by several vehicles. For the initial state of the network, the six distinct vehicle types were created in a route file, with the corresponding car-following model's parametrization. The route of each vehicle was also identified in the route file, as it was estimated from the path assignment of Aimsun.

In one hour of simulation for the morning peak, about 58% of the total demand was inserted in the network and 28% of the vehicles completed their journey within this time.

Table 4-9. Number of vehicles over the 10 replications of the simulation – Before recommendation

Replication	Seed number	Inserted vehicles	Active vehicles	Served vehicles
1	1042	49480	25258	24222
2	1043	49406	25706	23700
3	1044	48998	25558	23440
4	1045	50472	25791	24681
5	1046	47532	24849	22683
6	1047	50410	26041	24369
7	1048	49822	25612	24210
8	1049	50247	25717	24530
9	1050	50166	25519	24647
10	1051	48608	25191	23417
Average		**49,514 (58% of demand)**	**25,524**	**23,990 (28% of demand)**

Subsequently, for each vehicle that reached their destination the following parameters were estimated for each trip:

- average acceleration
- 90% percentile of acceleration
- average deceleration
- 90% percentile of deceleration
- speeding percentage

These driving characteristics were used as input to the RL controllers which recommend the optimal action for each trip. For the second scenario of simulation, the exact same vehicles were used, which follow the exact same routes on the same road network, in order to estimate the impact of the recommendation. The proposed actions of each vehicle were introduced as a modification of the car-following model's parameter in the route file. The adoption of this approach enabled hands-on implementation of the recommendation process with direct control over the outcomes. In this case as well, 10 replications with the same seed values as before, were performed to ensure the robustness of the results as presented in Table 4-10.

Table 4-10. Number of vehicles over the 10 replications of the simulation – After recommendation

Replication	Seed number	Inserted vehicles	Active vehicles	Served vehicles
1	1042	49052	25642	23410
2	1043	49449	25913	23536
3	1044	48375	25508	22867
4	1045	48092	25514	22578
5	1046	48704	25517	23187
6	1047	48153	25612	22541
7	1048	49192	25687	23505
8	1049	48421	25898	22523
9	1050	49964	26201	23763
10	1051	50470	25355	25115
Average		**48,987 (57% of demand)**	**25,685**	**23,302 (27% of demand)**

The detailed results of inserted, served and active vehicles revealed that in one hour of simulation 57% of the demand was served on average, while the corresponding percentage of served vehicles was reduced by 1% compared to the initial conditions.

4.4.3 Impact assessment through microsimulation

Any novel policy, technology or service has to be assessed before its large-scale implementation on the road network. In the case of Intelligent Transportation Systems (ITS) and according to the nature of each proposed system, the effects of its application can be evaluated in respect to various network aspects, that include traffic efficiency, road safety, social inclusion and pollution (Kaparias *et al.*, 2011). The literature demonstrates a detailed consideration of mobility and transport indicators which enables an efficient and straightforward monitoring of changes within a certain urban system.

The driving decision-support system proposed within this dissertation can produce personalized recommendations for any driver on the network and therefore, the impact of its application will be estimated through aggregated measures of traffic and pollution conditions, as well as road safety.

4.4.3.1 Overview and KPIs

Impact assessment of the proposed system is performed using microsimulation and by following a before-after approach. Specifically, for both simulation cycles the Key Performance Indicators of traffic, safety and environmental conditions were estimated, and comparatively assessed so that to quantify the overall impact of adopting personalized driving recommendations which improve each individual's driving behavior. The KPIs used in the analysis for each network's aspect are presented in Table 4-11.

Table 4-11. Key Performance Indicators for each network's aspect

Traffic	Safety	Environment
Served demand	Total conflicts	Cumulative amount of emissions (CO_2, CO, PMx, NOx)
MFDs	Total rear-end conflicts	Emissions per vehicle
	Conflicts per vehicle	

The estimation of traffic-related KPIs was dependent on the outputs of the simulation, which included the number of inserted and served vehicles, as well as edge-based information regarding the three fundamental elements of traffic flow theory (flow, speed and density). Instead of using aggregated measures of the fundamental variables, the Macroscopic Fundamental Diagrams were constructed and significant outcomes were drawn regarding the differences in the performance of the network before and after the application of the recommendation system.

The estimation of the harmful air pollutants is based on the emissions' model already integrated into SUMO, the PHEMlight model. PHEMlight is a simplified version of PHEM (Passenger car and Heavy-duty Emission Model), a complete vehicle emissions model developed in Europe since 1999 (Krajzewicz *et al.*, 2013, 2015). PHEM is based on extensive emission measurements on vehicles such as passenger cars, light duty vehicles and urban buses.

The approximation of the conflicts that constitute an indicator for road safety is based on the SSAM tool, which is presented in details in the next chapter.

4.4.3.2 Conflicts extraction with SSAM

The Surrogate Safety Assessment Model (SSAM) is a software application designed to perform statistical analysis of vehicle trajectory data output from microscopic traffic simulation models. The software computes a number of surrogate measures of safety for each conflict that is identified in the trajectory data and then computes summaries (mean, max, etc.) of each surrogate measure. A great illustration of a conflict detected between two vehicles is depicted in Figure 4-34.

Figure 4-34. An example conflict between two vehicles (Source: highways.dot.gov)

The most important measures of safety of each conflict are the time to collision (TTC) and the post encroachment time (PET). **TTC** refers to the minimum time-to-collision value observed during a conflict. This estimate is based on the current location, speed, and trajectory of two vehicles at a given moment. **PET** is the minimum post encroachment time observed during a conflict. Post encroachment time is the time between when the first vehicle last occupied a position and the second vehicle subsequently arrived at the same position. A value of 0 indicates an actual collision. Figure 4-35 presents the TTC and PET measures along with other safety performance measures of a conflict.

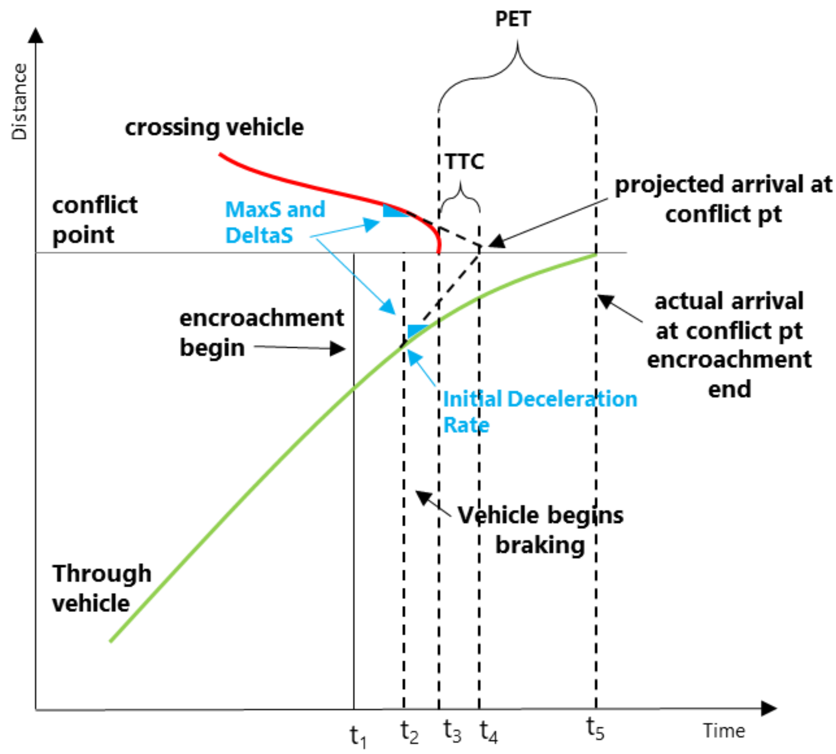


Figure 4-35. SSAM safety performance measures (Source: highways.dot.gov)

There are three type of conflicts that can be identified using the SSAM: crossings, rear-ends and lane changes. Conflicts' classification is mostly based on the conflict angle (θ), although link and lane information may be considered as well. A conflict angle is calculated for each pair of conflicting vehicles, based on the angle at which these vehicles converge to the potential collision point, as depicted in Figure 4-36.

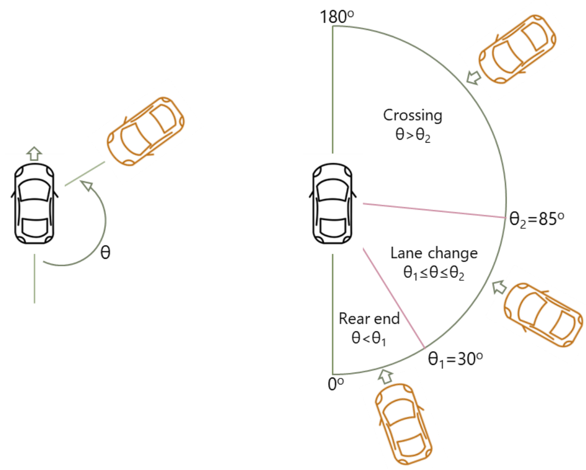


Figure 4-36. Conflict angle diagram

In case where two vehicles have a conflict on the same link and lane, then the conflict is a rear-end event regardless of the conflict angle. If both vehicles are on the same link and one of the vehicles changes lanes, then it is a lane-change event regardless of the conflict angle. However, link and lane information are not used if (a) it is not provided in the trajectory file, (b) the vehicles are on differing links, or (c) either vehicle changed links over the course of the conflict event. In such cases, the conflict angle is used for classification as follows:

- Rear end: $\|conflict\ angle\| < 30^o$
- Lane change: $30^o \leq \|conflict\ angle\| \leq 85^o$
- Crossing: $\|conflict\ angle\| > 85^o$

For the purposes of this research, the total number of conflicts are estimated as a Key Performance Indicator of road safety with the aim to evaluate the differences before and after the application of the recommendation system. A special focus is given on the rear-end conflicts, since the recommended driving actions refer to the car-following behavior of the driver, and no interventions are made into the lane changing model or the road infrastructure which are directly related with the lane changes and crossing conflicts respectively.

5 RESULTS AND DISCUSSION

In this chapter, results emerged from the analysis are described and explained. Initially, a comparison between the results produced by the two RL agents is provided. Then, the impact of applying the recommendation system is assessed based on the findings revealed from the two cycles of microsimulation. Impact assessment is performed in the basis of several Key Performance Indicators with regards to traffic, safety and the environment. Finally, a critical discussion on the main results of this dissertation is provided.

5.1 Driving recommendations

The ultimate goal of this dissertation was to develop an inclusive methodological framework that incorporates two main functionalities: the identification of driving behavior and the provision of personalized driving recommendations. Advanced Deep Reinforcement Learning agents were trained for this purpose, enabling the generation of driving recommendations with respect to two categories of drivers; typical drivers who exhibit a moderate average behavior and unsafe drivers who interchange their behavior among various unsafe driving habits.

The fitted models were used to produce recommendations for some indicative trips as shown in Table 5-1. The RL controllers take as input the characteristics of a trip (average acceleration and deceleration, 90% quartile of acceleration and 10% of deceleration, percentage of trip with speed over the speed limit) and propose either an increase or a decrease in the maximum and the average acceleration. The produced alterations refer to the optimal driving actions that the specific driver can adopt based on their current behavior. In some cases, an increase in average or maximum acceleration or even both may be suggested, which may be due to a variety of facts:

1. All the variables describing driving behavior are inextricably linked with each other and therefore, an increase in the acceleration value may involve the improvement of another critical driving parameter. In such a case, even the adoption of a higher acceleration may be a safer decision for the driver.
2. According to the rationale of Reinforcement Learning algorithms, a proposed action corresponds to the best possible action that the specific driver could perform instantly, so that in the future they may progress into improved safer behaviors, given the way they are currently driving.

Even from this small sample of trips, it is evident that both the RL controllers are trained to generate recommendations that move drivers closer to the average safe behavior of a typical driver, which has an average acceleration equal to 1.137 m/s^2 and a maximum acceleration equal to 2.503 m/s^2. For the indicative samples of Table 5-1, the mean recommended average acceleration was estimated 1.145 m/s^2, while the mean value of the proposed maximum accelerations was 2.507 m/s^2 respectively. It can therefore be concluded that a universal application of the proposed recommendation system would lead to the harmonization of the acceleration profiles for the entire fleet of vehicles.

In the sections that follow, a discussion of the results of each version of the RL algorithm is provided based on targeted examples of trips and their corresponding recommendations.

Table 5-1. Example of RL input and output and the produced recommendations

Profiling and Scoring		RL input					RL output		Recommendation	
Driving profile	Score	Acceleration average	Deceleration average	Acceleration Q90	Deceleration Q90	Speeding percentage	action_1 (max)	action_2 (avg)	new Acceleration Q90	new Acceleration average
Aggressive-distracted	9	1.512	-1.58	3.600	-3.72	0.0090	↓ -0.99	↓ -0.31	2.610	1.197
Aggressive-distracted	2	1.888	-1.86	3.336	-3.95	0.0660	↓ -0.29	↓ -0.51	3.046	1.377
Aggressive-risky	6	1.779	-1.63	3.432	-3.82	0.3235	↓ -0.65	↓ -0.49	2.778	1.288
Aggressive-risky	2	1.574	-1.73	3.480	-4.20	0.5693	↓ -0.96	↓ -0.40	2.519	1.171
Aggressive	2	1.905	-1.68	3.600	-3.18	0.0000	↓ -0.77	↓ -0.63	2.834	1.275
Aggressive	17	1.948	-2.57	3.768	-5.69	0.0822	↓ -1.10	↓ -0.66	2.669	1.289
Aggressive	10	1.774	-2.17	4.368	-4.08	0.0000	↓ -1.69	↓ -0.54	2.676	1.230
Distracted	2	1.405	-1.33	4.104	-2.88	0.0133	↓ -1.38	↓ -0.21	2.727	1.193
Distracted	24	1.504	-1.52	3.324	-3.36	0.0000	↓ -0.64	↓ -0.27	2.679	1.236
Distracted	35	1.054	-1.08	2.489	-2.67	0.0000	↑ 0.06	↑ 0.11	2.548	1.166
Moderate	73	1.386	-1.59	2.760	-3.28	0.0226	↓ -0.46	↓ -0.33	2.299	1.055
Moderate	58	0.592	-0.92	1.560	-2.81	0.0294	↑ 0.56	↑ 0.39	2.123	0.980
Moderate	86	1.366	-1.43	2.880	-3.36	0.0000	↓ -0.45	↓ -0.26	2.426	1.106
Moderate	90	0.826	-1.03	1.984	-2.43	0.0000	↑ 0.26	↑ 0.18	2.240	1.006
Moderate	67	0.967	-1.43	2.040	-3.67	0.0674	↑ 0.05	↓ -0.01	2.091	0.954
Moderate	78	1.047	-1.23	2.070	-2.33	0.0000	↑ 0.21	↓ -0.02	2.284	1.030
Moderate	77	0.666	-0.69	1.440	-1.68	0.0367	↑ 0.76	↑ 0.32	2.195	0.989
Moderate	94	1.124	-1.24	2.568	-2.64	0.0000	↓ -0.06	↓ -0.01	2.511	1.118
Moderate	58	0.871	-1.24	1.920	-3.56	0.1667	↑ 0.27	↑ 0.13	2.187	1.003
Risky	52	1.367	-1.36	3.093	-3.06	0.2392	↓ -0.49	↓ -0.17	2.602	1.196
Risky	33	1.098	-1.43	2.280	-3.24	0.2857	↑ 0.32	↑ 0.09	2.600	1.192

5.1.1 Typical drivers' controller – SADRA I

Personalization is an integral aspect of the proposed system as it is a key feature for all recommendation systems, since it can lead to higher acceptance rates and better user experience, and consequently to users' long-term engagement with the system. For this reason, in this dissertation, special focus was given to the development of distinct controllers that produce personalized recommendations by matching users' previous driving behavior to either of the two versions of the RL agent. Here, the outcomes of SADRA I are presented, which is the version of the RL agent that generates recommendations that correspond to a typical/moderate driving behavior. The RL controller can suggest either an increase or a decrease in the average and the maximum acceleration for the next trip of each driver. Indicative results for each of the dimension of the action vector are presented in Figure 5-1 and Figure 5-2 which depict alterations in the average and maximum acceleration respectively, for a number of random trips.

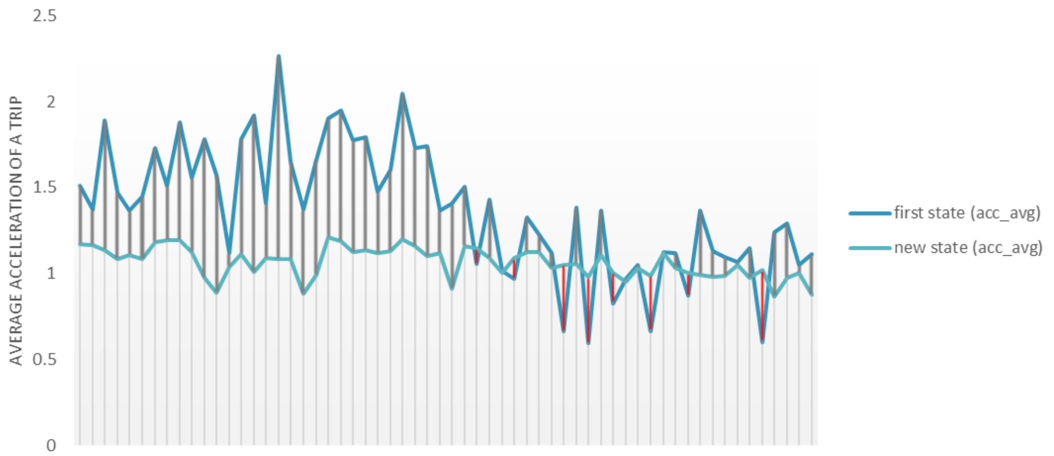


Figure 5-1. The difference between the first and the new state's average acceleration after the recommendation produced from the RL typical controller. The increase of the average acceleration is depicted with red color, while the decrease is depicted with grey color.

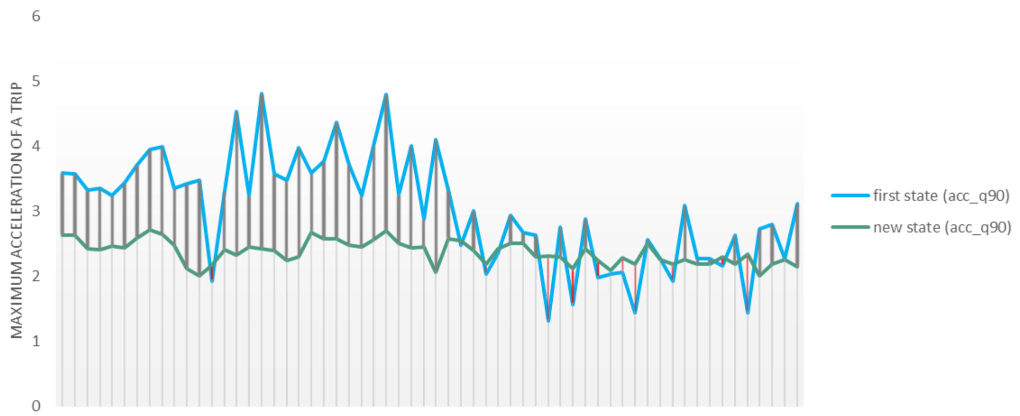


Figure 5-2. The difference between the first and the new state's maximum acceleration after the recommendation produced from the RL typical controller. The increase of the maximum acceleration is depicted with red color, while the decrease is depicted with grey color.

It is evident that all drivers are guided towards the average safe behavior which emerged from the clustering procedure, but yet, every driver receives a different recommendation that perfectly matches their own driving behavior. Therefore, a driver who had a very low value of maximum acceleration on their trip, they would be recommended to increase its value on their next trip.

5.1.2 Unsafe drivers' controller – SADRA II

The second version of the RL agent produces recommendations for drivers who exhibit an overall unsafe driving behavior during their trips. The application of the RL controller to a number of driving trips performed by drivers who drive in a more aggressive and unsafe manner, resulted in significantly safer accelerations.

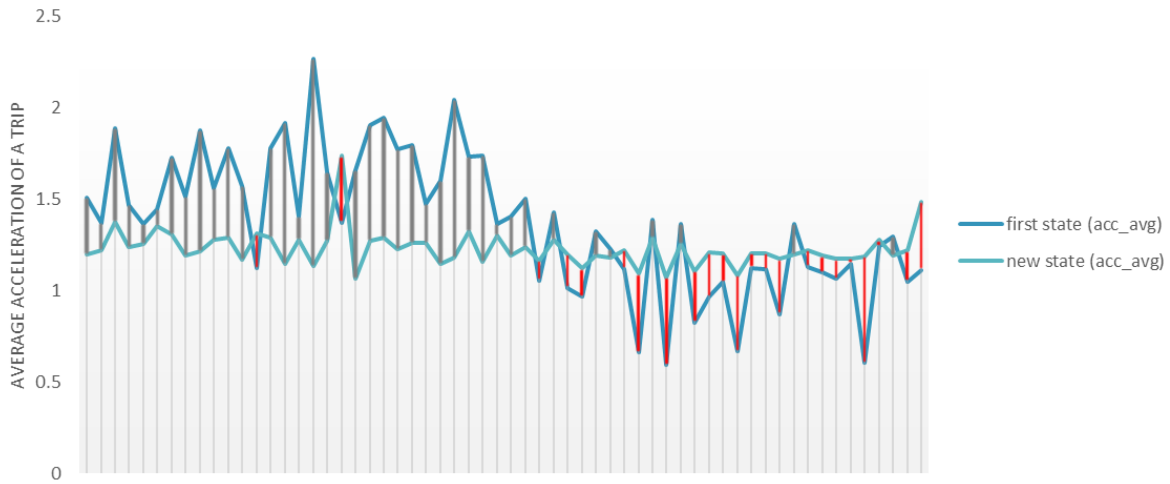


Figure 5-3. The difference between the first and the new state's average acceleration after the recommendation produced from the RL unsafe controller. The increase of the average acceleration is depicted with red color, while the decrease is depicted with grey color.

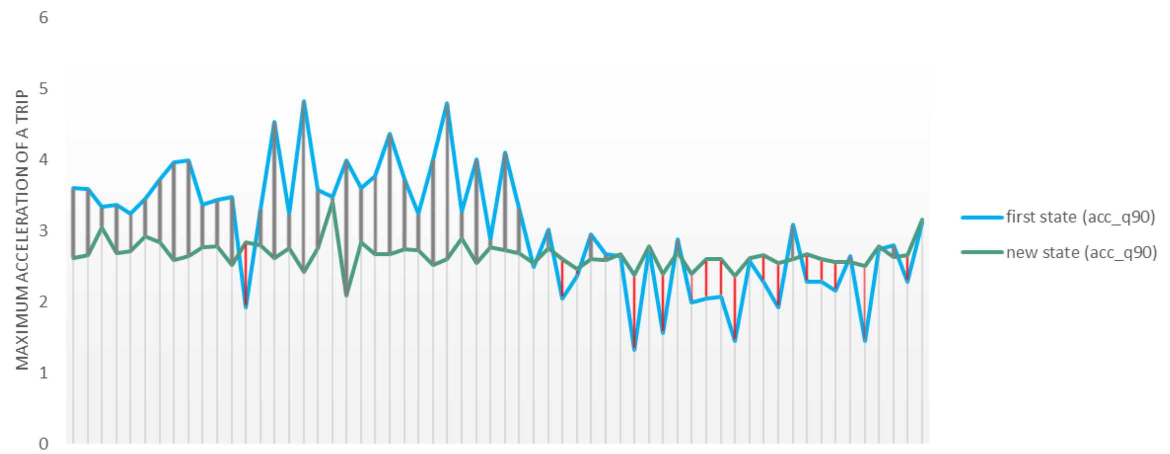


Figure 5-4. The difference between the first and the new state's maximum acceleration after the recommendation produced from the RL unsafe controller. The increase of the maximum acceleration is depicted with red color, while the decrease is depicted with grey color.

5.1.3 Output comparison of the controllers

In this section the output of the two controllers is comparatively discussed. Through indicative examples of driving trips, the differences between the produced recommendations are presented, with the aim to highlight the importance of personalization as a core aspect of the proposed system.

The interpretation of the results, as they are depicted in Figure 5-5 and Figure 5-6, indicates that the two controllers lead to different driving states. Although the recommendations of the controller concerning unsafe drivers (SADRA II) lead to significantly lower average accelerations for the next trip (next state) compared to the previous trip (initial state), they maintain a significant distance upwards for the respective recommendations produced from the typical drivers' RL controller (SADRA I). Nevertheless, one should notice that both the controllers lead to a smoother acceleration profile for the entire traffic.

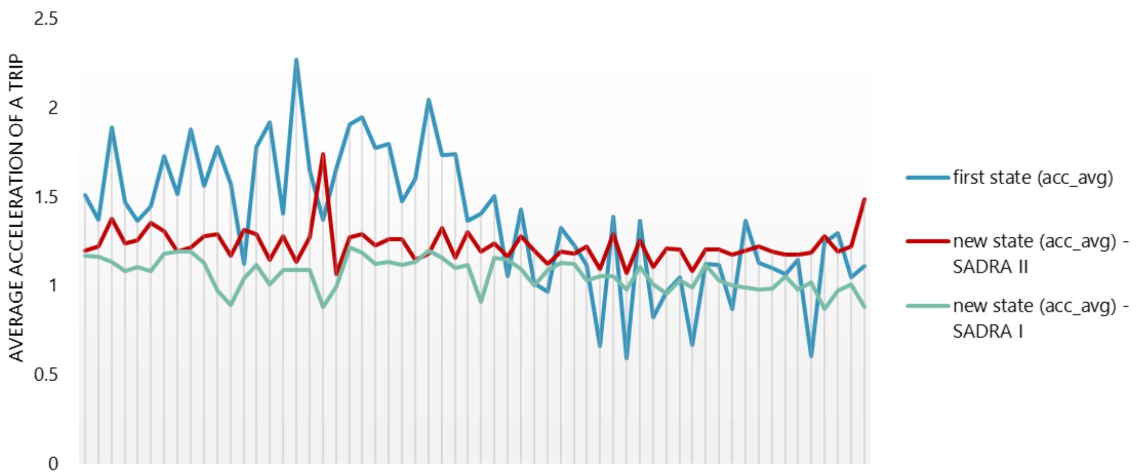


Figure 5-5. Comparison of the new state's average acceleration as it emerged from the Typical and the Unsafe RL controllers.

The results concerning the maximum accelerations of the corresponding trips present a similar picture to that of the average accelerations. As before, recommendations produced from both controllers lead to a smoother profile though the SADRA II suggests greater maximum acceleration values than those produced from the controller of typical drivers (SADRA I).

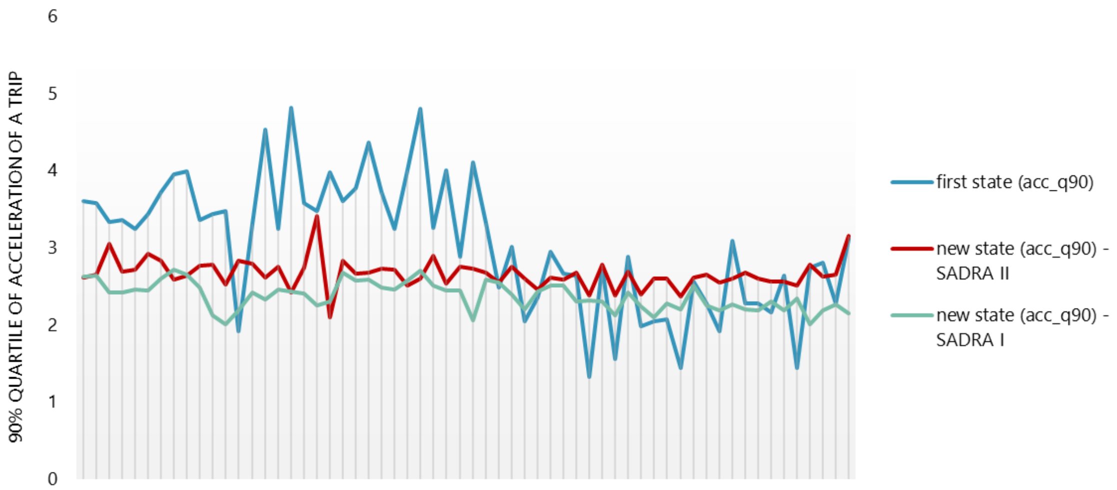


Figure 5-6. Comparison of the new state's maximum acceleration as it emerged from the Typical and the Unsafe RL controllers.

These results are quite as expected, as each controller was trained using sequences of trips from drivers of the respective profiles (typical/unsafe) and, thus, it provides actions that are compatible with the corresponding driving styles. Furthermore, it is also what we aimed at, as a, for example, aggressive driver would not follow a recommendation for very smooth acceleration that would be very different from what they are used to.

5.2 Application of the recommendation system

The quantification of the impact of applying the proposed recommendation system and in consequence, of the adoption of an improved driving behavior by all drivers is of great importance both for researchers as well as practitioners and can lead to significant findings regarding the usefulness of improving individual driving behavior. The assessment of the recommendation system is performed by utilizing specific Key Performance Indicators that correspond to three areas of interest: traffic, safety and emissions.

In this section, the results of applying the recommendation system will be presented and discussed, as emerged from a network-level traffic simulation using the calibrated Athens Road Network. Two rounds of the simulation scenario were performed corresponding to the initial state of the road network and the conditions after the application of the recommendations, respectively. Each of the simulation rounds was done in 10 replications to enhance the validity and robustness of the results. In total, the trained SADRA I controller was used to produce recommendations for 43% of the vehicles, while the rest of the vehicles followed the recommendations produced by SADRA II. Figure 5-7 graphically represents the number of inserted and served vehicles for both simulation rounds.

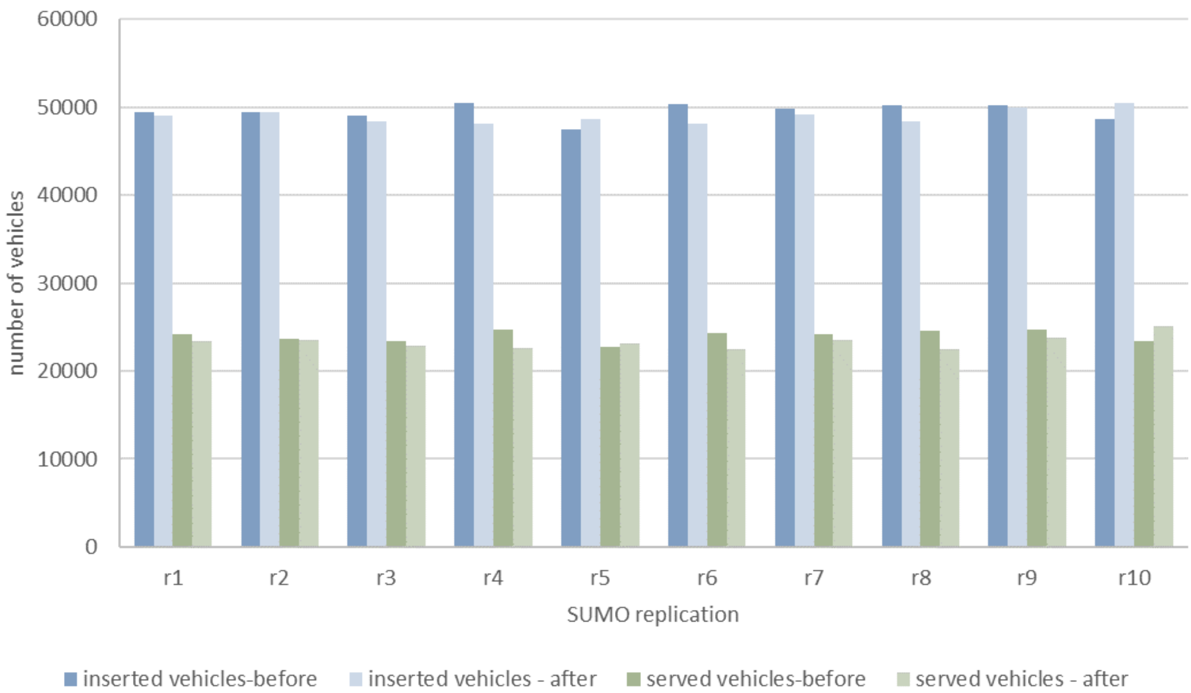


Figure 5-7. The number of served and active vehicles over one hour of simulation for the 10 replications

All replications of the same simulation round present mutual results which are slightly reduced after the application of the recommendation system. A first glimpse on the results revealed that when vehicles follow the proposed recommendations the performance of the road network may be to some degree decreased, in the sense that less vehicles reach their destination within the one hour of simulation. On average, 2.9% less vehicles were served based on the results of the second round of the simulation.

However, results of the statistical hypothesis test *t-test* indicated that there are no significant differences between the means of the served vehicles before and after the recommendations in 95% confidence interval. As shown in Table 5-2, the alternative hypothesis is rejected.

Table 5-2 Paired t-test results for served vehicles

	mean	std	t-value	p-value
Served vehicles – before	23989.9	655.7	1.845	0.098
Served vehicles – after	23302.5	783.1		

A greater investigation of the traffic flow properties together with aggregated metrics of driving behavior will be conducted to further quantify the potential impact on the other dimensions of the road network as well. The application of the personalized recommendation system had a substantial impact on the maximum acceleration of the vehicles, as shown in Figure 5-8. When all vehicles followed the suggestions generated by the two RL controllers, the mean value of the maximum acceleration was somewhat increased from 2.83 m/s^2 to 2.96 m/s^2, mostly because the majority of the vehicles who adopted a very small maximum acceleration, which was far lower from the corresponding acceleration of the "moderate/typical" behavior, they were suggested to slightly increase their acceleration. However, the condensation of the interquartile range is evident after the recommendations, which indicates the harmonization of the acceleration profiles of all vehicles in the simulation. Finally, the maximum value of the observed maximum accelerations remained at the same level of 3.94 m/s^2 after the application of the proposed system.

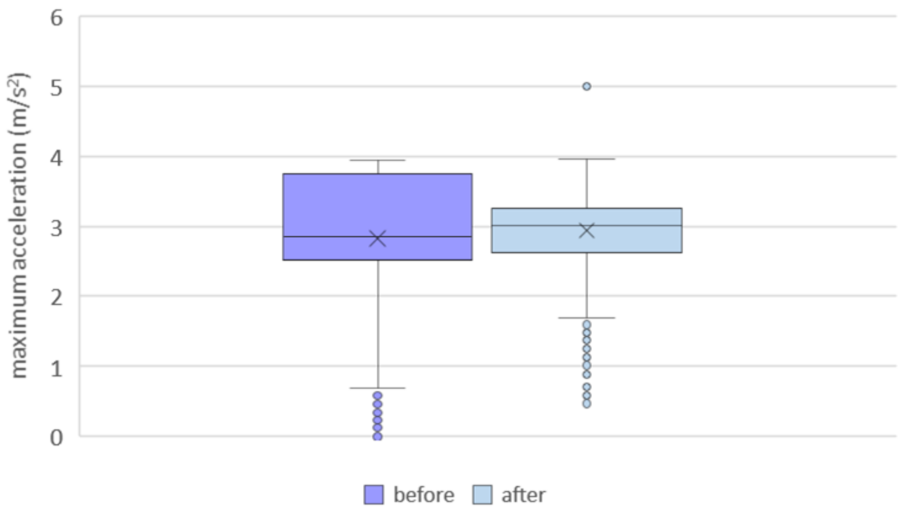


Figure 5-8. Boxplot of maximum acceleration before and after recommendation

All the variables that describe driving behavior are inextricably linked with each other and therefore, the changes that were imposed on the acceleration of the vehicles after the recommendations, led to changes in the rest driving behavior parameters as well. As far as it concerns the average and the maximum speed performed by all vehicles in the simulation, several critical findings can be revealed when comparing the statistical characteristics of the two variables before and after applying the recommendations. Specifically, on the one hand, the differences observed in the magnitude of the average speed are minimal, since in both situations the vehicles adopt an average speed of around 25 km/h, while the maximum average speed that is observed is approximately 55 km/h. A summary of the statistic of the average speed is nicely depicted in Figure 5-9.

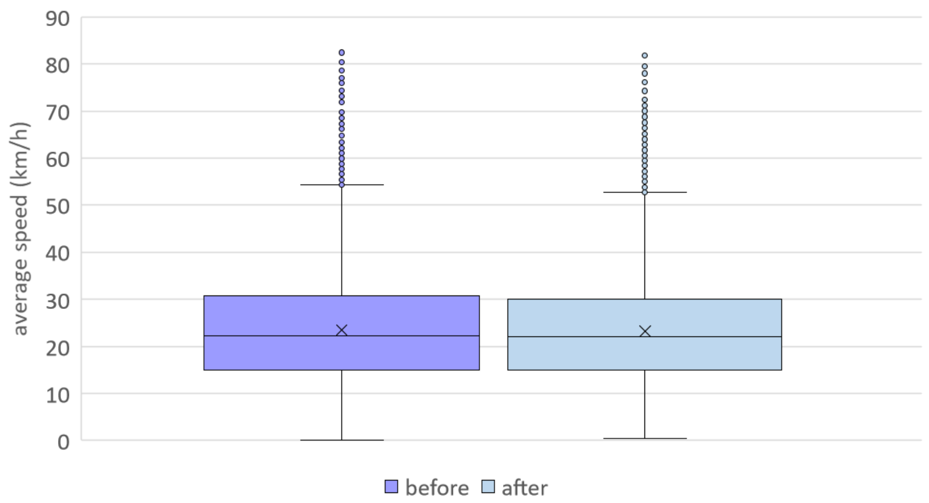


Figure 5-9. Boxplot of average speed before and after recommendation

On the other hand, a noteworthy difference is evident mostly on the maximum values of speed where the majority of vehicles' speed is concentrated around the mean value and the interquartile range is pointedly narrower than the corresponding values of the initial conditions of the network (Figure 5-10). Such findings can be thought as anticipated considering the fact that the proposed recommendation system eliminates the extreme accelerations and generally smooths out the accelerations for the entire traffic. It can be concluded that the harmonization of the maximum accelerations led to a corresponding normalization of the maximum speeds.

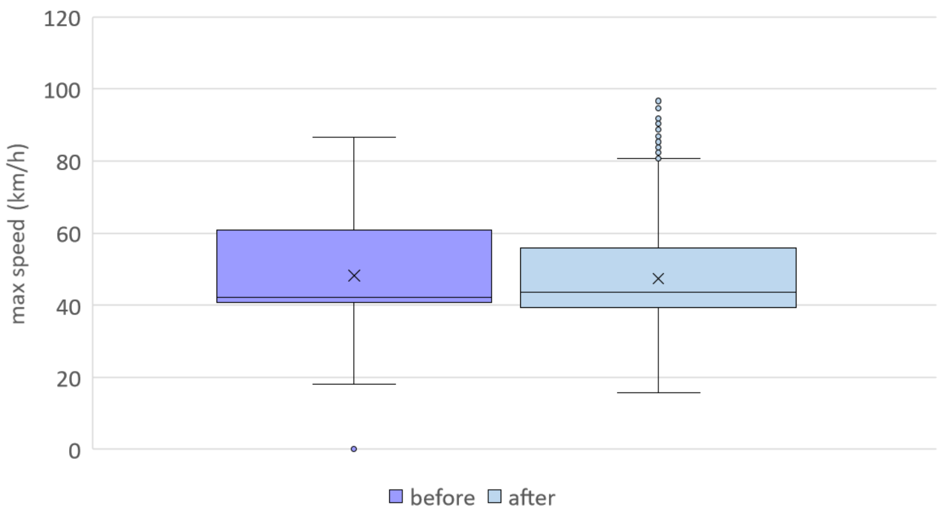


Figure 5-10. Boxplot of maximum speed before and after recommendation

Alterations on the speed of vehicles resulted on changes of the rest traffic flows properties, namely flow and density. Microscopic fundamental diagrams were calculated to provide a thorough graphical representation of these variables' relations for the initial conditions as well as the conditions emerged after the recommendations. All three fundamental diagrams (Figure 5-13, Figure 5-12, Figure 5-13) demonstrate the relationships between traffic flow properties, namely mean vehicle flow, mean density and mean speed, as they emerged from the simulation based on aggregated measurements of all edges for the 10 replications. Results indicate that the implementation of self-aware driving suggestions although it leads to safer and less aggressive driving behavior for each individual, it does not improve the performance of the road network. More specifically, self-improvement is evident from the lower mean density values which indicates that vehicles keep greater distances from the leading vehicles. Additionally, lower speeds are also observed after the adaptation of the recommended accelerations with the difference from the initial conditions being more significant in the case of saturated network flow (Figure 5-13).

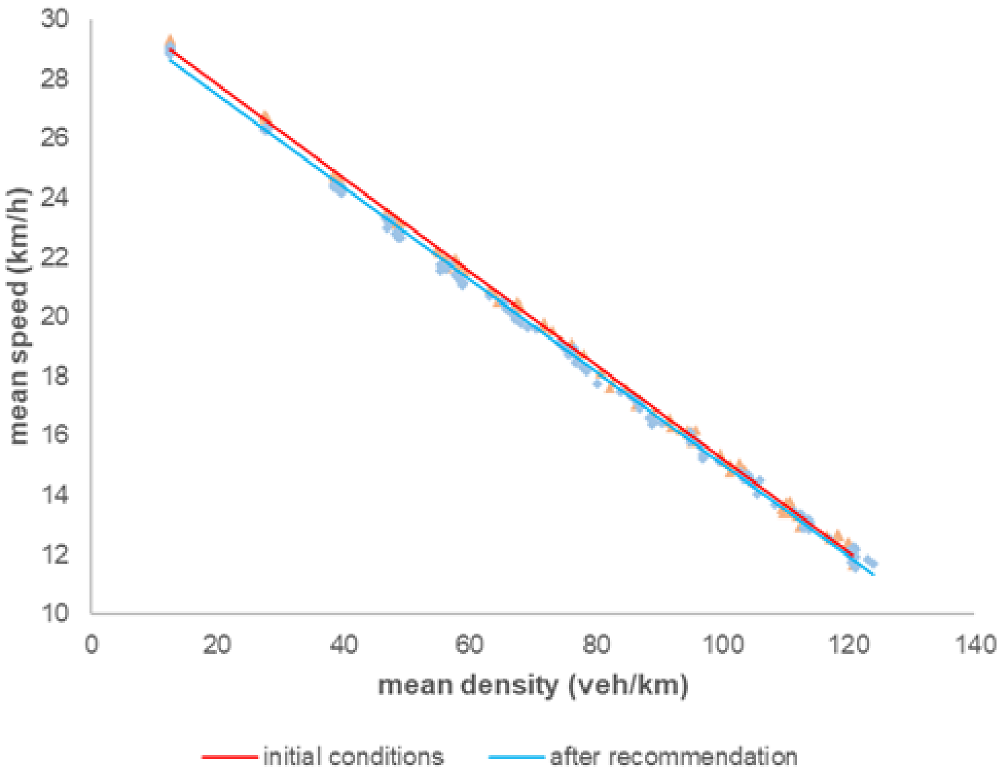


Figure 5-11. Fundamental diagram of speed-density before and after driving recommendations, based on simulation results

Individual driving safety is augmented, yet the impact on traffic conditions is not similarly positive. The vehicles that move at lower speeds and with a lower density worsen traffic flow conditions, since fewer vehicles are served per time unit compared to the initial conditions. Nonetheless, this decrement of mean flow may be considered acceptable if assessed in conjunction with the positive effects on driving safety. However, based on the findings of this research, it can no way be concluded that the improvement of personal driving behavior is associated with a significant improvement in traffic conditions and therefore, the imposition of soft policy measures, such raising self-awareness with respect to individual driving safety and performance, it cannot be considered as a key measure for traffic management.

The fundamental diagram of flow-density seems to depict a uniformity between the initial and the final conditions, although some minor differences are observed with respect to the absolute value of capacity flow. Specifically, for the value of critical density, which was estimated 33.1 veh/km, the corresponding values of traffic flow are 360 veh/h and 358 veh/h for the initial conditions and after the recommendations respectively.

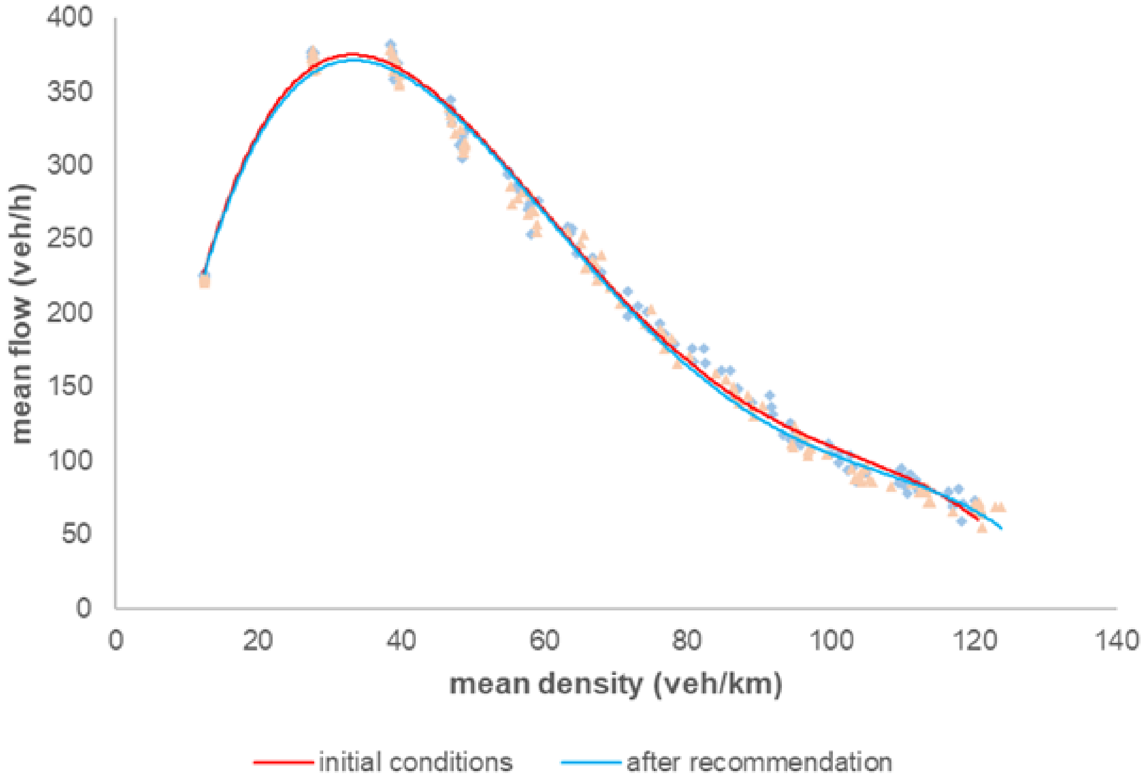


Figure 5-12. Fundamental diagram of flow-density before and after driving recommendations, based on simulation results

The flow-speed diagram is used to determine the speed at which the optimum flow occurs. For the initial conditions of the road network, the optimum flow occurs when vehicles move with 26.1 km/h, while the corresponding speed after the recommendation is reduced 3.4% with its absolute value estimated 25.2 km/h.

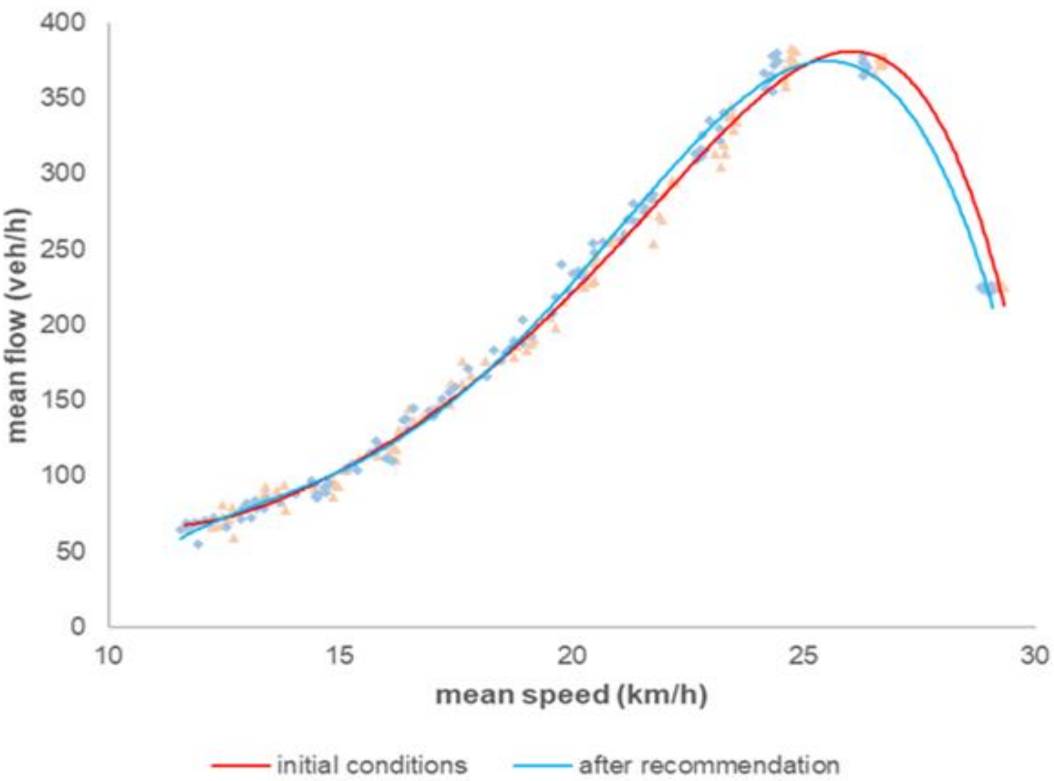


Figure 5-13. Fundamental diagram of flow-speed before and after driving recommendations, based on simulation results

Except for the performance of the network, another key performance indicator is safety. Improved driving safety was evident from lower values of density which indicated that vehicles kept greater distances from the leading cars, and additionally, from the lower observed mean speed. The assessment of the applied recommendations with respect to safety was performed by calculating the number of conflicts occurred between the vehicles during the simulation. The detection of the critical conflicts was performed using the SSAM tool which was fed with the trajectories of the vehicles as they emerged from the two cycles of simulation. Table 5-3 presents the number of conflicts that were observed for the entire traffic before and after the recommendation. There are three types of conflicts that can be identified from vehicles' trajectories, which are crossings, rear-ends and lane changes. Here, a special focus on rear-ends is given since the proposed recommendations only affect the car-following behavior of each driver. Results presented in Table 5-3 indicate that the application of a driving recommendation system significantly improves road safety.

Table 5-3. Safety performance indicators in Athens Network before and after applying driving recommendations

	Initial conditions	After recommendation [% difference]
Vehicles served	23,990	23,302
(in one hour of simulation)	*(27.88% of demand)*	*(27.08% of demand)*
Total number of conflicts	2.86 conflicts/vehicle	2.75 conflicts/vehicle [-4.2%]
Rear - ends	2.01 rear-ends/vehicle	1.90 rear-ends/vehicle [-5.5%]

A reduction of 4.2% of the total number of conflicts was observed when vehicles followed the corresponding driving recommendations, while the corresponding percentage of elimination for the rear-end conflicts is 5.5%. Although these percentages may not seem very high, the absolute number of conflicts that was calculated after the recommendation is significantly reduced by approximately 6,000 conflicts for the one hour of simulation. Rear-ends constitute about 33% of the total number of conflicts, which indicates that each driver gets involved in all different kind of conflicts during driving. The detailed table with the results for both before and after the implementation of the system of recommendations conditions is given in Table 5-4.

Table 5-4. Number of conflicts before and after the recommendations

Rep.	BEFORE			AFTER		
	Total number of conflicts	Rear-ends	Conflicts/vehicle	Total number of conflicts	Rear-ends	Conflicts/vehicle
1	130169	44111	2.63	129314	43168	2.64
2	149844	48800	3.03	132268	43481	2.67
3	142947	47774	2.92	130606	41701	2.70
4	134756	46084	2.67	128406	44226	2.67
5	156078	53508	3.28	147744	47135	3.03
6	124487	42273	2.47	118456	41663	2.46
7	146748	50553	2.95	144624	46605	2.94
8	138355	47460	2.75	129545	42240	2.68
9	144159	49698	2.87	135286	44931	2.71
10	147855	51174	3.04	149530	47614	2.96

In the case of conflicts, the statistical hypothesis test *t-test* indicated that there is a significant difference between the means of both the total conflicts and the rear-ends before and after the recommendations. As presented in Table 5-5, the alternative hypothesis is accepted in a 99% confidence interval, especially for the rear-ends conflicts which is the main KPI for the specific research in terms of the impact on road safety.

Table 5-5 Paired t-test results for conflicts

	mean	std	t-value	p-value
Total conflicts – before	141539.8	9617.2	3.898	0.004
Total conflicts – after	134577.9	9834.7		
Rear ends – before	48143.5	3369.5	5.860	<0.001
Rear ends – after	44276.4	2227.0		

Some indicative results on the impact of the proposed recommendation system on emissions is provided. The corresponding Key Performance Indicator is the level of emissions for all different kind of air pollutants, namely Carbon Dioxide (CO_2), Carbon Monoxide (CO), Particulate Matter (PM_x) and Oxides of Nitrogen (NO_x). A significant reduction in all categories of emissions is observed compared to the initial conditions of the network, as shown in Table 5-6. Findings revealed that the homogenization of acceleration profile for the entire traffic has led to a slightly reduced emissions per vehicle. Specifically, the reduction in all categories of emissions is estimated as follows: 2.5% in CO_2, 0.3% in CO, 1.3% in PM_x and 3.3% in NO_x. It should be noted that this improvement in the environmental conditions is very important since the proposed recommendation system had a positive impact on emissions despite the fact that the controller was not trained towards this direction.

Table 5-6. Difference in vehicle emissions before and after applying driving recommendations

Emissions	Initial conditions	After recommendation [% difference]
CO_2	0.704 kg/vehicle	0.686 kg/vehicle [-2.5%]
CO	0.027 kg/vehicle	0.026 kg/vehicle [-0.3%]
PM_x	0.0133 g/vehicle	0.0131 g/vehicle [-1.3%]
NO_x	0.296 g/vehicle	0.287 g/vehicle [-3.3%]

To sum up, the application of a system that provides personalized recommendations for improved driving on a network-wide level leads to the harmonization of the acceleration profile for the entire traffic. The adoption of uniform accelerations from the vehicle fleet resulted in the occurrence of lower mean and maximum speeds in the network. It was also observed that when vehicles followed the proposed recommendations and they accordingly adapted their acceleration, greater distances were kept from the leading vehicles, which led to in lower mean density values. For these reasons, slightly less vehicles were served in the network for the same period of simulation time. Therefore, it can be concluded that the proposed system improves the performance of each driver individually without leading to improvements in the traffic flow conditions for the network. Instead, **traffic conditions** seem to get slightly worse after the system has been implemented.

Contrary to the impact on traffic, the proposed recommendation system has a significant positive impact on driving and **road safety**. Driving safety is enhanced since each individual driver is being recommended the best action that they can perform based on their current driving state. Improvements in road safety are evident from the lower rate of conflicts that correspond to each vehicle when the initial conditions of the network are compared to those after the recommendation. Specifically, the harmonization of the acceleration profile of the fleet resulted in a reduction of 4.2% of the total number of conflicts, while the corresponding percentage for the rear-end conflicts was 5.5%. The significant reduction of the rear-ends is of great importance since such conflicts are the result of the driver's car-following behavior in which interventions are made through the implementation of the recommendation system.

Finally, a noteworthy lessening in the level of **emissions** for all different kind of air pollutants is observed when vehicles follow the proposed recommendation. It should be highlighted once again that although results regarding the emissions are presented here for the sake of completeness, they are not further discussed in detail since no interventions was made by the proposed system in the emission model, on the contrary, their estimation was based solely on the default model available in SUMO simulation software.

5.3 Main results and discussion

The summary of the results shaped by this research includes key points in the following axes:

- Driving profiling
- Model development
- Driving recommendations
- Impact assessment

First, using a two-level clustering approach, it was feasible to identify and distinguish between aggressive and non-aggressive driving behavior, and subsequently, to detect six driving behaviors that a driver can exhibit during their trip. Spanning from the typical – safe behavior, to aggressive driving and unsafe – reckless driving, the entire range of driving behavior was detected and the six driving profiles were used to annotate the behavior of each driving trip. Further investigation of driving behavior on a user level has highlighted the existence of two main driver categories; the first one includes drivers that mostly perform a moderate driving behavior free from aggressiveness and reckless driving, while the other one refers to drivers who mostly exhibit several unsafe behaviors during driving.

Then, it was shown how Deep Reinforcement Learning algorithms can be exploited in order to determine optimal policies for each individual driver so that to improve their driving behavior. It should be noted that this is the first time that naturalistic driving data are used within the framework of RL algorithms with the aim to produce actions for real driving. The models developed were based on the Actor-Critic RL approach, where two neural networks were trained; the first one estimates the Q-values for all possible actions and the other one matches the states with the corresponding best action. All neural networks converged after a number of epochs and were able to produce driving recommendations that improve driving behavior both in the short- and long- term.

Two distinct RL controllers were trained, each corresponding to a specific group of drivers, so that to match suggested driving actions to individual driving preferences. Findings revealed that for the same driving state the two agents produce different recommendations. Specifically, both controllers lead to the harmonization of the acceleration profile of all drivers, yet the controller corresponding to unsafe drivers proposes slightly greater accelerations compared to the controller of the typical drivers.

Impact assessment of the proposed system is performed through a microsimulation setting for the Athens' Road Network using SUMO simulation software. Results indicate that the implementation of self-aware driving suggestions although it leads to safer and less aggressive driving behavior for each individual, it does not lead to improved traffic conditions. Specifically, after the recommendation, vehicles move at lower speeds and road segments are occupied with a lower density and therefore, slightly less vehicles were served over the one hour of simulation. However, one may claim that the deteriorating traffic conditions can be considered acceptable if one takes into account the compensation through the benefits of adopting smoother driving behavior in road safety and driving comfort. Specifically, using the number of conflicts as a key performance indicator of road safety, it is observed that the adoption of the recommended actions leads to a significant improvement in the total number of conflicts as well as refining car-following behavior of each vehicle. In addition to enhancing road safety, the homogenization of the acceleration ability of the vehicles also leads to slightly reduced emissions of air pollutants. In line with numerous previous studies, the adoption of smoother acceleration profile from the entire traffic eliminates harmful emissions and improves environmental conditions.

6 CONCLUSIONS

6.1 Overview

Driving behavior has been in the spotlight of research for a variety of reasons ranging from understanding the dynamics of driving behavior through innovative concepts, such as crowd-sensing, to developing self-improvement frameworks and building human-like behavioral models for autonomous vehicles. The ability of researchers to identify driving behavior and most importantly to detect unsafe driving habits has given an impetus to the development of driving assistance and scoring systems. Nevertheless, the already developed systems lack the wide acceptance and establishment since they cannot adapt to each drivers' personal preferences and needs. Within this context, the ultimate goal of this dissertation was to develop a novel driving recommendation framework which aims at improving individual behavior in terms of driving aggressiveness and riskiness based on a data-driven methodological approach. This ultimate goal encloses specific objectives that should be achieved sequentially, which answer the research questions put forward after an exhaustive review of the literature.

Initially, this research addressed key research questions concerning driving behavior dynamics and specifically, it identified the main driving profiles that describe decision making while driving. An inclusive methodological framework was implemented in order to extract driving profiles straight from the smartphone crowd-sensed data using unsupervised learning. Moreover, the level of aggregation of the overall driving behavior is investigated with the aim to understand the extent to which driving behavior can be categorized in groups that reflect different driving styles. This dissertation, also, provided answers to the critical questions of whether Artificial Intelligence can be exploited to resemble human decision making especially in the complex task of driving and further, to select the most appropriate Reinforcement Learning algorithm for supporting driving decisions. Furthermore, this research attempted to provide answers to the critical question of whether there is a link between raising self-awareness and improving conditions of the entire network. Finally, answers are given with rewards to the impact of controlling individual driving behavior on driving and road safety, as well as to the environmental conditions.

An inclusive methodological framework was proposed to achieve the objectives of this dissertation, which incorporates a mixture of statistical analysis, machine learning techniques and reinforcement learning algorithms. The plethora of tools and methods used, enabled the understanding of driving behavior dynamics through smartphone data, the development of the personalized recommendation system and the assessment of its impact through a large-scale microsimulation scenario. Specifically, starting from smartphone crowd-sensed data interesting driving features were extracted and exploited in a two-level k-means clustering approach. The first level of clustering distinguishes aggressive from non-aggressive trips, while the second level of clustering resulted in the identification specific driving profiles.

Subsequently, the driving footprint of each driver was estimated as the average behavior of all their trips. Users included in the dataset were separated in groups based on their driving footprint so that drivers in the same groups share common driving characteristics.

The development of the recommendation system was based on the training of RL agents each one corresponding to a specific group of drivers with common driving behavior. In this way, the agents were able to produce recommendations that explicitly match each drivers' preferences and therefore, increase the possibilities of being adopted by them. In order to train the RL controllers, their main components had first to be defined, namely states, actions, and rewards. Based on the nature of the problem (studying of driving behavior), both states and actions are continuous and therefore, the actor-critic structure was chosen as the most appropriate RL algorithm. The reward function was also constructed from scratch with the aim to rewarding safe/typical driving behavior while penalizing unsafe driving habits (distraction, risk taking) and by exploiting a custom-made score function which assigns a score to each trip according to the level of driving safety. Model development was based on two neural networks whose hyperparameters emerged after an exhaustive grid search.

Once the RL agents were trained, their performance was assessed through simulation, and more precisely, using the Athens' Road Network, a large-scale simulation scenario was set aiming at quantifying the impact of applying the personalized recommendation system. Impact assessment was performed on the basis of traffic, safety and emissions, and thus, proper Key Performance Indicators were defined. A before-after approach was followed to calculate the impact on these three dimensions, where the initial conditions of the network ("before") were simulated using the characteristics of the emerged driving profiles as distinct vehicle types. Then, the RL agents produced personalized recommendations for each individual vehicle, and a second run of simulation was performed where all vehicles followed the corresponding driving actions ('after"). The performance of the network before and after the application of the system was assessed through Macroscopic Fundamental Diagrams. As far as it concerns driving and road safety, they were evaluated through measurements of speed, acceleration and distances from the leading cars, as well as critical conflicts between the vehicles.

The proposed methodological framework led to the production of critical conclusions in all dimensions of the phenomenon under consideration. The main findings can be summarized in the following points:

- A two-level clustering approach can provide great insights on the characteristics that govern aggressiveness during driving and can be further exploited to distinguish safe from unsafe driving patterns.
- Six distinct driving profiles are able to describe the overall driving behavior that someone performs during their trip.
- There are two categories of drivers according to the average behavior of each driver resulting from how they drove in all their trips. In the first category drivers usually drive

in a typical manner while in the second category drivers perform a number of unsafe driving actions or drive in an aggressive manner in the majority of their trips.

- The Actor-critic approach from the family of reinforcement learning algorithms can be exploited to find the best possible driving action for each dividual driver given the way they drove in their previous trip.
- When a controller provides driving recommendations to a fleet of vehicles, the acceleration profile of the entire fleet is harmonized on a value which is close enough to the acceleration decisions of a typical – safe driver.
- The application of a personalized recommendation system to a city's road network does not have a significant impact on traffic conditions. In contrast, it is shown that slightly less vehicles can be served for the exact same simulation period.
- When each driver improves their own behavior, road safety is enhanced on the network. Specifically, critical conflicts between vehicles are significantly reduced after the application of the proposed system.
- The level of emissions for all different kinds of air pollutants is reduced which indicates that harmonization of the accelerations for the entire traffic can have an important positive impact on the environmental conditions.

Concluding, it should be noted that the deterioration of traffic may be considered acceptable if one takes into account the compensation through the benefits of adopting smoother driving behavior in road safety and emissions. To this end, policy makers and researchers should not neglect the real impact on all network's dimensions when planning traffic management strategies and applying soft and hard policy measures. In the following sections, the way each stakeholder could benefit from the findings of this work are discussed.

6.2 Main contributions

The present doctoral dissertation offers significant innovative contributions in five areas:

It makes use of an innovative naturalistic driving dataset. A great volume of data was available with high temporal resolution from real driving, enriched with a variety of factors that describe driving behavior, environment and other external attributes for each trip.

It proposes a methodological framework to extract driving profiles straight from the data, which describe the entire range of driving behavior. A data-driven approach is followed to classify critical driving patterns that appear during a trip by exploiting k-means clustering as being the most appropriate tool for this purpose.

It develops novel Reinforcement Learning algorithms to solve a real-world problem, this of assisting driving behavior. A deep Reinforcement Learning algorithm was chosen as the most suitable tool to learn the optimal policy and suggest the appropriate action that leads to the best possible driving behavior for each individual driver.

It proposes a methodology which is capable of recognizing individual driving preferences and produce personalized driving actions to each driver. Specifically, an inclusive methodological framework is implemented which incorporates tools and methods that first recognize driving behavior of every user, then assigns every user to the corresponding RL controller version based on their overall behavior and finally produces personalized driving actions that mitigate aggressiveness and riskiness of driving.

It evaluates the large-scale network effects of implementing a personalized driving recommendation system on three areas of interest using specific KPIs, precisely on traffic, safety and emissions. Impact assessment of the proposed recommendation system is performed using a real-world scenario that of the Athens' Road network through microsimulation and by applying a before-after methodology to compare the values of the KPIs before and after the application of the system.

This research contributes to the scientific field of driving behavior analytics through big data in both a conceptual, technical and a practical level. First of all, through an extensive review of the literature, it sheds light to all critical challenges and caveats that both research and practitioners may face when collecting and analyzing data collected through smartphone crowd-sensing. Despite the fact that crowd-sourcing has become very popular over the last decades as an effective and cost-efficient source of data, most researchers neglect the critical issues that arise when collecting data from the crowd. This dissertation not only records these challenges and discusses ways to address them, but through the methodological framework it applies it goes beyond most of them in the development of the proposed driving recommendation system.

Equally important is the contribution of the dissertation to the conceptual design of a data-driven methodology that is able to detect a variety of human behaviors and choices during driving. Specifically, in this research we have shown that the behavior a driver exhibits during a trip can be characterized by either safe or unsafe maneuvers. Findings revealed that the entire range of behaviors on a trip level can be described by six driving profiles spanning from safe driving to aggressiveness, risk taking and distraction during driving. In addition, it was shown that the above driving behaviors can be identified directly from data gathered through smartphone crowd-sensing by following an unsupervised learning approach. To the best of the authors' knowledge, no effort has been made previously to identify normal and abnormal driving patterns without following a threshold-based methodology, but instead relying fully on a data-driven approach. This research succeeds in detecting driving behavior on a trip level using a two-level k-means clustering algorithm on a variety of driving parameters that explicitly describe how a driver behaved during a specific trip. Moreover, a significant contribution of this research is that it highlighted aggressiveness during driving as a distinct driving style that can be observed both as the unique driving behavior characteristic, but also it can be detected simultaneously with the occurrence of other unsafe driving behaviors such as speeding and distraction. The implementation of the k-means clustering algorithm at two distinct levels enabled the identification of aggressive driving as a unique driving characteristic that defines a driver's choice of driving style independently to the rest driving actions (safe or unsafe) performed during a trip.

Contrary to already developed ADAS and other recommendation systems that impose flat driving suggestions or follow rule-based approaches, the proposed framework automatically formulates personalized quantified driving actions, to be disseminated to drivers as recommendations for future behavior without posing unrealistic restrictions on personal driving style. Therefore, for the same driving behavior during an initial trip the implementation of the developed RL controllers based on the overall behavior of the specific driver would lead to different recommendations. This way, a more aggressive driver is more smoothly driven into the transition to calmer driving characteristics (driving with reduced maximum and average accelerations over their trip) compared to a typical driver. The personalization aspect of the proposed system is vital, since it is linked with users' acceptance and engagement rates. However, relevant works have neglected the importance of self-awareness when developing driving assistance systems, in the sense that the system is aware of personal characteristics and preferences when producing the corresponding recommendations and the driver becomes aware of the impact that their driving choices have on individual driving performance and safety as well as the network and the environment. The developed self-aware driving recommendation system can be exploited to assist drivers and provide them with feedback on their total driving efficiency in a variety of concepts such as recommendation systems, ADAS, insurance telematics and many more. An additional value of the methodology proposed is that it can be hard implemented as is in most of the aforementioned platforms. Moreover, the flexible architecture of the algorithms used to design the proposed recommendation system enables its application even with a different overall scope only by applying some modification on the reward function. For example, although the ultimate goal of the system is the improvement of individual driving behavior, meaning that each driver receives the optimal recommendation that leads him to an improve driving state, a change on the reward function can cause the system to provide recommendations for eco-driving, meaning that each driver will receive the optimal recommendation that leads him to more efficient fuel consumption during their trip.

Furthermore, this work contributes to the exploitation of advanced reinforcement learning algorithms for solving a real-world problem, that of developing a decision support system for drivers. Reinforcement Learning is not a new concept, but only recently it gained a lot of attention mostly due to the advent of big data. Nevertheless, most research focuses on exploiting RL algorithms on solving low-dimensional problems such as games and robots' movements. In this work, we have used advanced RL algorithms to develop a recommendation system for drivers that is capable of producing continuous actions for real driving, and for this purpose we have adjusted accordingly the Deep Deterministic Policy Gradient algorithm in a way that it can proposes the best possible driving alterations in each given state. The proposed approach and implementation are on-of-a-kind both due to the large-scale naturalistic driving dataset that is exploited as well as due to the complexity of the phenomenon in terms of states and actions. This is the first time that the DDPG algorithm is adjusted and implemented within a recommendation framework for real-driving while at the same time being trained using naturalistic driving data. This research contributes to the developing field of Reinforcement Learning and provides significant insights regarding the implementation of theoretical algorithms to real-world problems.

On the top of these, this study makes use of innovative data that were collected using a smartphone application, which is based on state-of-the-art algorithms for the recording, collection and processing of crowd-sensed data. Smartphone crowd-sensing is an approach that is becoming popular nowadays and is considered as a cost-efficient and effective solution for collecting data in naturalistic driving experiments. The main contribution of this research with regards to big data handling relies mostly on the way that high-dimension data were exploited within the methodological framework without increasing computational complexity while at the same time ensuring data quality and analysis accuracy. More specifically, following a step-wise methodological approach it was possible to manage of large volumes of data in an efficient way since data dimensionality is reduced based on feature selection procedures and data are separated to different subsets with similar characteristics.

Finally, another contribution of this work is the application of the proposed system to a real road network through microsimulation. Since scientific community has not determine the impact of personalized driving assistance systems, part of this dissertation was devoted to the analysis and quantification of this impact through targeted KPIs. Impact assessment is performed on the basis of three areas of interest, namely traffic, safety and emissions. Contrary to previous research which had assumed that when each individual improves their driving behavior and moves carefully and safely the entire network would benefit from improved traffic conditions, in this work we explicitly studied this impact using specific KPIs. Findings revealed that traffic flow conditions do not benefit from the large-scale implementation of such a system, despite the fact that road safety is significantly enhanced. These results give a great impetus to further research and the investigation of other theories such as the price of anarchy and information dissemination strategies, which are discussed later on.

6.3 Research limitations

Limitations of this study with regards to the data and each part of the methodology adopted, are briefly presented in this section. If these limitations are addressed, the proposed recommendation system can gain significant prospects as a tool for managing traffic and enhancing road safety.

As any other data-driven approach, this research as well, relied on some limitations with regards to problem setup and adaptation. Firstly, some limitations emerged from the need to match the RL output with the simulation properties. More specifically, one of the two components of the recommended action, the average acceleration of each driver, could not be imported into the microsimulation environment since only the maximum acceleration is taken into consideration and the instantaneous acceleration is defined by a variety of factors (such as leading vehicle's speed etc.) and, thus, limits the ability of evaluating the entire potential of the proposed recommendation system. The car-following model adopted by the microsimulation platform used in this dissertation, is parametrized by a plethora of factors governing vehicle's dynamics such as velocity, acceleration and deceleration. The Krauss model is parametrized by the maximum acceleration of the vehicle which corresponds to the acceleration ability of vehicles of a specific type, and therefore only the one component of the recommended action is adopted within the simulation (Figure 6-1). Nevertheless, due to the nature of the phenomenon of driving, all parameters describing how a driver chooses to drive over a trip are inextricably linked with each other and therefore, the neglection of the average acceleration was not expected to have a significant impact over the results of the simulation.

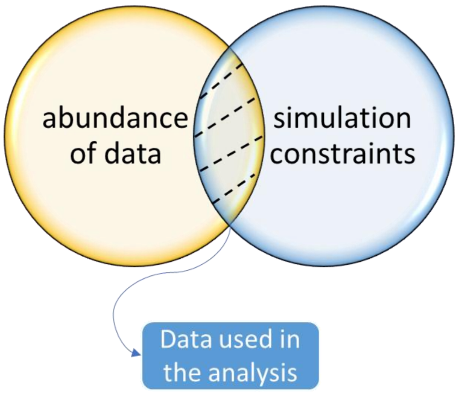


Figure 6-1. Schematic representation of the criteria used for data selection

Besides car-following behavior, a driver during their trips takes actions regarding lane change, priority concession and other decisions concerning interactions with other road users. However, in this research the focus was explicitly on the car-following behavior as the ultimate goal was to create a user-centric system that looks only at the driver and does not require any external information from the road network in order to be trained and implemented. Thus, the proposed actions refer on the way the driver drives along the road, namely the way they choose to hit the acceleration pedal, which depends only on the personal preferences and perceptions of the driver. The provision of recommendations regarding lane change behavior or any other behavior that the driver performs, e.g., the way they decelerate or the adaptation of vehicle speed, would have required the availability of exogenous data including, driving environment and surroundings, traffic, road geometry etc., and therefore the system would cease to be autonomous, but instead it should have been context-aware. However, the lack of information about the environment can be considered as a limitation of the developed system, since its transformation into a context-aware system would give other perspectives both to the system itself and to the possibilities of its use as a traffic management tool.

An extension of the above limitation is the fact that since the system ignores the state of the environment it cannot operate real-time. In other words, the proposed methodology is not able to produce recommendations real-time, namely during a trip. Instead, an offline system is developed which suggests alterations on driving behavior in a sequence of trips for each driver. The integration of external information into the system would allow, at least conceptually, the real-time provision of driving recommendations. However, even in this case, data availability remains a caveat for the development of such a system, since it would require the existence of high-resolution data that could be recorded, stored and analyzed on the fly within the framework of an online recommendation system. Moreover, the real time operation of the proposed system would have required modifications on the main components of the RL model. More specifically, in the case where the system becomes context-aware and operates in real-time, the state of the environment should include parameters that describe traffic and road conditions, the proposed actions should be adapted to the restrictions imposed by the road, traffic signaling and the rest of the traffic, and finally, the reward function should incorporate the trade-offs between the improvement of individual driving and the impact on traffic, safety and emissions.

Lastly, another limitation, which applies to all data driven approaches, is the generalization and transferability of the developed model and the corresponding outcomes. In most cases it is unclear whether the sample used to train the model is representative of the entire population and also whether its characteristics are similar to those of a different population. In this work, a big naturalistic driving dataset is used to develop the RL models which includes trips performed by a great number of drivers, nevertheless, it cannot be said that the results can be generalized and spatially transferred to another road network. When human behavior is being examined, it should be borne in mind that in addition to environmental constraints, available options and prevailing conditions, model transferability may be hindered by other factors associated with cultural and ethical differences among individuals. From a more technical point of view, one particular limitation of this study is the tuning of the hyperparameters for the artificial neural networks as these were selected based on a specific dataset. As a result, if a new dataset is applied to the trained models, the hyperparameters may need to be recalibrated. Finally, another potential limitation of these techniques is that deep learning methods can be considered a 'black box' method and therefore might lack interpretability for different stakeholders and traffic managers. However, the deep RL models developed within this dissertation can be easily assessed and interpreted straight through the produced output which corresponds to recommendations for actual driving behavior.

6.4 Research impact

This doctoral dissertation develops a novel driving recommendation system using artificial intelligence to enhance driving safety of each individual driver by providing recommendations that improve driving performance without neglecting personal preferences of each driver. Large amounts of data were collected by an already developed application through smartphone crowd-sensing and thus, a variety of driving parameters were collected and further exploited within this work.

In order to fulfill its goal, which is to create a self-aware system of providing personalized driving recommendations, this thesis developed algorithms for detecting driving behavior and then developed, trained and applied advanced deep reinforcement learning algorithms to produce the appropriate recommendations for each driver according to their driving conditions. Currently, there are limited regulations on validating the impact of these systems even though evaluating the outputs from such systems are important. Therefore, this research contributes by implementing the developed novel recommendation system in a virtual network using an integrated simulation framework. The outputs produced within this dissertation may have a significant impact on several aspects of both research (R), technology (T) and policy-making (P):

(R) Aggressiveness does not necessarily constitute an unsafe driving habit and can be detected either as an individual behavior or in combination with other unsafe behaviors.

This research exploited a large-amount of naturalistic driving data, during the collection of which the drivers did not receive any information about the way they were driving nor any guidance or other information that might affect their behavior and decisions during driving. This fact gave to our research an extra perspective as it allowed the investigation of the whole range of driving behaviors that includes both extreme unsafe behaviors and more restrained, typical driving behaviors. Findings revealed that a driver can exhibit a certain driving behavior over their trips which is either an aggressive or a non-aggressive behavior. On the top of this decision, the driver may choose to either perform additional unsafe driving maneuvers namely distraction or speeding, or just choose to travel only in an aggressive manner or even perform none of these behaviors and thus drive safely. Such results can have a great impact on the field of driving analytics and pave the way for studying driving aggressiveness as an independent driving behavior which may provide answers to the existing question "Does aggressiveness constitute an unsafe driving habit?"

(R) Reinforcement learning algorithms can be implemented in real-world problems and specifically, the DDPG algorithm can learn how to make human-like decisions on complex and high-dimensional environments.

Until now, Reinforcement Learning algorithms have been widely used to learn games and in robotics, and, thus, in this dissertation we took on the challenge of applying a deep reinforcement learning algorithm to train an agent to make decisions just as the driver would. To the best of the author's knowledge, this is the first time that the Deep Deterministic Policy Gradient algorithm is successfully implemented within the framework of a real-world problem that of improving driving behavior and relying explicitly on real data (not simulation or synthetic data). A great number of researchers can benefit from the conceptual design and custom setup of the DDPG algorithm performed within this work with the aim to produce recommendations for human driving control. More specifically, the fact that the proposed methodology is fully transferable and interpretable enables the adoption of the proposed approach in a multitude of problems that require the modelling and controlling of human decision making (continuous actions) in complex multidimensional environments (continuous states). For example, the developed system can be implemented with the aim to promote eco-driving or in other words, to recommend those actions that can improve fuel efficiency. In addition, it can be implemented in other transport-related domains, such as autonomous vehicle's operation, traffic signalization and so on.

(T) The proposed recommendation system can be incorporated in already developed software, such as insurance telematics apps and ADAS.

Besides the innovation aspects of the proposed methodology and the novel algorithm developed, this dissertation contributes to technology as well, since it developed a read-to-implement platform which incorporated two main functionalities; i) it can identify the driving profile of each driver per trip and estimate every driver's overall driving footprint and ii) it can massively produce personalized driving recommendations that improve individual driving performance. The system can be useful as part of Usage-Based services, namely pricing schemes based on driving usage or characteristics i.e., Pay-How-You-Drive driving insurance schemes.

Moreover, the developed recommendation system can be incorporated in an ADAS framework, which aims to support driving behavior in a more personalized way by adjusting actions to the preferred driving style of each driver, rather than implementing predefined generalized actions. The recommendation system developed within this dissertation can revolutionize the assistance provision system and pave the way for the new-generation ADAS.

(R & T) The identified human driving profiles can provide great insights for human-like autonomous driving.

Driving uniformity is an important factor in road safety, and therefore, a significant challenge of autonomous driving is to imitate, while remaining within safety bounds, human driving styles, or in other words to achieve human-like driving. The results outlined in this research both from the driving behavior analysis and the driving recommendation provision can have far reaching implications in the development of state-of-the-art behavioral models for autonomous cars. First of all, the six identified profiles together with their driving characteristics can provide great insights on the different types of drivers that coexist on the road, and therefore, facilitate the design of surroundings detection and comprehension systems of autonomous vehicles, as well as improve the interaction protocols under mixed traffic conditions. In addition, the understanding of differences on the driving dynamics between the different driving profiles can enable the development of more human-like behavioral models, in the sense that autonomous vehicle control can be performed either in a very conservative and cautious way or in a more bold and aggressive manner always complying with safety requirements. A great body of literature deals with the development of human-like driving models for self-driving cars for two main reasons:

a) Human-like driving will enhance the confidence of passengers while riding an autonomous car
b) Human-like behaving cars will facilitate the understanding of surrounding drivers about the movements of the autonomous cars as well as the interaction with conventional cars and other road users.

(P) The improvement of driving behavior on an individual level can have significant impact on road safety and emissions, but not a noteworthy impact on traffic conditions.

It is a common assumption of researchers that the improvement of driving behavior in an individual level can result in an improvement of traffic. Nevertheless, none of the previous studies have examined the actual impact of improving each driver's behavior on the entire traffic. Therefore, in this research special focusing was given to the quantification and assessment of individual driving recommendation provision on traffic, safety and emissions. Our findings revealed that although the improvement of driving behavior on an individual level can have significant impact on road safety and emissions, as expected, the corresponding impact on traffic conditions is not noteworthy. Such findings are of great importance both for researchers and policy-makers since current assumptions have to be reconsidered and traffic management strategies that highlight driving improvement as a key factor for traffic conditions enhancement have to be updated.

6.5 Future research directions

Within this dissertation we developed a self-aware driving recommendation system, using a mixture of unsupervised learning and reinforcement learning algorithms and by exploiting an innovative naturalistic driving dataset, with the aim to improve driving behavior through the mitigation of aggressiveness and riskiness. The impact of the provision of personalized driving recommendations is assessed through a city-wide microsimulation scenario by properly adjusting traffic models. Findings of this work can have far reaching implications for future research.

The analysis of driving behavior occupied much of this dissertation and resulted in the definition of driving behavior on two levels: on a trip level where 6 distinct driving profiles were detected and, on a user-level, where the driving footprint of each driver is estimated based on their overall driving behavior. However, there is still much room in the exploration of driving behavior dynamics and thus, further research should be conducted in that direction involving enriched driving datasets and additional driving behaviors and parameters (e.g., cornering, tasks that cause distraction except from mobile usage). Moreover, the dedicated study of the dynamic evolution of driving behavior is also very important to provide answers to the question of how much and how rapidly driving profiles are altering over time. It is known that individual driving behavior is characterized by large values of volatility, but under which conditions a driver adopts a different driving style and how differently can a driver drive from trip to trip?

Another direction of future research concerns the recommendation system. Initially, the already developed recommendation system can be adjusted properly to produce driving recommendations with a different goal. For example, it would be of great interest to produce driving actions with the goal to achieve fuel efficiency during a trip and in addition, to compare these driving recommendations with the ones produced with the aim to improve driving safety. How different are the driving behavior that should be adopted from the assisted drivers in these two cases?

Another important research question raised at this point refers to the communication between the user (driver) and the system. Future research should investigate the way the produced recommendation should be passed to the driver in order for them to be understood by the user and then to be accepted by him. Although this interaction between the user and the system is a key success factor of any service, it usually being neglected during the design process of the service. Especially in the case of human driving assistance, the produced recommendations have to be clear enough for the user to adopt them and therefore, future research should definitely study this aspect as well.

A critical challenge of driving recommendation system is the real-time operation, which has not been addressed within the framework of this dissertation, due to the limitations discussed in the previous section. The identification of the required specifications that will enable the real-time operation of the system could also be a part of future research. Towards this direction, the most significant future research objective would be the modification of the proposed system in a way that it becomes context-aware, meaning that the system can interact with environment in which the agent takes decisions and have a full view of its dynamics and alterations. In this way, the proposed system could be implemented in real-time, and additionally it could also act as a traffic management tool which uses driving behavior as a key force of enhancing traffic efficiency.

The proposed methodological approach based on a mixture of unsupervised learning and reinforcement learning strategies can be valuable for the development of easily adaptable behavioral models especially for partially automated vehicles. Such models' prerequisites, except from driver personality and styles identification, also include situation awareness and behavioral adaptation. For this reason, future work can focus on building the context for the RL agents to be placed in, namely become context-aware, and, additionally, enrich the RL components with parameters describing traffic conditions and road geometry.

REFERENCES

Abdulazim, T., Abdelgawad, H., Nurul Habib, K. and Abdulhai, B. (2013) 'Using Smartphones and Sensor Technologies to Automate Collection of Travel Data', *Transportation Research Record: Journal of the Transportation Research Board*, (2383), pp. 44–52. doi: 10.3141/2383-06.

Abou Elassad, Z. E., Mousannif, H., Al Moatassime, H. and Karkouch, A. (2020) 'The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review', *Engineering Applications of Artificial Intelligence*, 87(August 2019), p. 103312. doi: 10.1016/j.engappai.2019.103312.

Afsar, M. M., Crump, T. and Far, B. (2021) 'Reinforcement learning based recommender systems: A survey'. Available at: http://arxiv.org/abs/2101.06286.

Amini, S., Tilg, G. and Busch, F. (2018) 'Evaluating the impact of real-time traffic control measures on the resilience of urban road networks', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2018-Novem, pp. 519–524. doi: 10.1109/ITSC.2018.8569678.

Andrieu, C. and Pierre, G. Saint (2012) 'Using statistical models to characterize eco-driving style with an aggregated indicator', *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 63–68. doi: 10.1109/IVS.2012.6232197.

Angkititrakul, P., Ryuta, T., Wakita, T., Takeda, K., Miyajima, C. and Suzuki, T. (2009) 'Evaluation of Driver-Behavior Models in Real-World Car-Following Task', *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, November, pp. 113–118.

Araújo, R., Igreja, Â., De Castro, R. and Araújo, R. E. (2012) 'Driving coach: A smartphone application to evaluate driving efficient patterns', *IEEE Intelligent Vehicles Symposium, Proceedings*, 1(1), pp. 1005–1010. doi: 10.1109/IVS.2012.6232304.

Bamberg, S., Fujii, S., Friman, M. and Gärling, T. (2011) 'Behaviour theory and soft transport policy measures', *Transport Policy*, 18(1), pp. 228–235. doi: 10.1016/j.tranpol.2010.08.006.

Barnett, V. and Lewis, T. (1994) *Outliers in Statistical Data*. John Wiley.

Bejani, M. M. and Ghatee, M. (2018) 'A context aware system for driving style evaluation by an ensemble learning on smartphone sensors data', *Transportation Research Part C: Emerging Technologies*, 89(February), pp. 303–320. doi: 10.1016/j.trc.2018.02.009.

Bergasa, L. M., Almeria, D., Almazan, J., Yebes, J. J. and Arroyo, R. (2014) 'DriveSafe: An app for alerting inattentive drivers and scoring driving behaviors', in *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 240–245. doi: 10.1109/IVS.2014.6856461.

Bhattacharya, R. and Waymire, E. C. (2007) *A Basic Course in Probability Theory*. New York: Springer, ISBN: 978-3-319-47974-3.

Bhoraskar, R., Vankadhara, N., Raman, B. and Kulkarni, P. (2012) 'Wolverine: Traffic and road condition estimation using smartphone sensors', in *2012 4th International Conference on*

Communication Systems and Networks, COMSNETS 2012. IEEE, pp. 1–6. doi: 10.1109/COMSNETS.2012.6151382.

Bieker-Walz, L., Behrisch, M., Junghans, M. and Gimm, K. (2017) 'Evaluation of car-following-models at controlled intersections', *31st Annual European Simulation and Modelling Conference 2017, ESM 2017*, (Krauss), pp. 238–243.

Birenboim, A. and Shoval, N. (2015) 'Mobility research in the age of the smartphone', *Annals of the American Association of Geographers*, 106(2), pp. 283–291. doi: 10.1080/00045608.2015.1100058.

Bishop, C. (2006) *Pattern Recognition and Machine Learning*, p. 738, ISBN: 978-0-387-31073-2.

Bonsall, P., Liu, R. and Young, W. (2005) 'Modelling safety-related driving behaviour - Impact of parameter values', *Transportation Research Part A: Policy and Practice*, 39(5), pp. 425–444. doi: 10.1016/j.tra.2005.02.002.

Bora, D. J. and Gupta, A. K. (2014) 'A comparative study between fuzzy clusterign algorithm and hard clustering algorithm', *International Journal of Computer Trends and Technology (IJCTT)*, 10(2).

Bose, B., Dutta, J., Pramanick, P., Roy, S. and Ghosh, S. (2018) 'Smartphone based system for real-time aggressive driving detection and marking rash driving-prone areas', pp. 1–6. doi: 10.1145/3170521.3170549.

Bricka, S., Zmud, J., Wolf, J. and Freedman, J. (2009) 'Household Travel Surveys with GPS', *Transportation Research Record: Journal of the Transportation Research Board*, 2105, pp. 51–56. doi: 10.3141/2105-07.

Butakov, V. and Ioannou, P. (2015) 'Personalized Driver/Vehicle Lane Change Models for ADAS', *IEEE Transactions on Vehicular Technology*, 64(10), pp. 4422–4431. doi: 10.1109/TVT.2014.2369522.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbride, A. and Goodwin, P. (2008) 'Smarter choices: Assessing the potential to achieve traffic reduction using "Soft measures"', *Transport Reviews*, 28(5), pp. 593–618. doi: 10.1080/01441640801892504.

Cao, H. and Lin, M. (2017) 'Mining smartphone data for app usage prediction and recommendations: A survey', *Pervasive and Mobile Computing*, 37, pp. 1–22. doi: 10.1016/j.pmcj.2017.01.007.

Castanedo, F. (2013) 'A review of data fusion techniques', *The Scientific World Journal*, 2013. doi: 10.1155/2013/704504.

Castignani, G., Derrmann, T., Frank, R. and Engel, T. (2015) 'Driver behavior profiling using smartphones: A low-cost platform for driver monitoring', *IEEE Intelligent Transportation Systems Magazine*, 7(1), pp. 91–102. doi: 10.1109/MITS.2014.2328673.

Chakravarty, T., Ghose, A., Bhaumik, C. and Chowdhury, A. (2013) 'MobiDriveScore–A System for Mobile Sensor Based Driving Analysis', in *2013 Seventh International Conference on Sensing*

Technology, pp. 338–344. Available at: http://seat.massey.ac.nz/conferences/icst2013/proceedings/papers%5C1569809327.pdf%5Cn http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6727671.

Chan, T. K., Chin, C. S., Chen, H. and Zhong, X. (2019) 'A Comprehensive Review of Driver Behavior Analysis Utilizing Smartphones', *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–32. doi: 10.1109/tits.2019.2940481.

Chang, Y.-J., Paruthi, G., Wu, H.-Y., Lin, H.-Y. and Newman, M. W. (2016) 'An investigation of using mobile and situated crowdsourcing to collect annotated travel activity data in real-word settings', *International Journal of Human Computer Studies*, (November), pp. 1–22. doi: 10.1016/j.ijhcs.2016.11.001.

Chatzimilioudis, G., Konstantinidis, A., Laoudias, C. and Zeinalipour-yazti, D. (2012) 'Crowdsourcing with Smartphones', *IEEE Internet Computing*, 16(5), pp. 36–44.

Chen, Z., Yu, J., Zhu, Y., Chen, Y. and Li, M. (2015) 'D3: Abnormal driving behaviors detection and identification using smartphone sensors', *2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking, SECON 2015*, pp. 524–532. doi: 10.1109/SAHCN.2015.7338354.

Chiu, Y.-C., Ma, Y.-L. and Hu, X. (2014) 'Identifying Driving Risk Factors to Support Usage-Based Insurance using Smartphone Personalized Driving Data', *Workshops on Big Data and Urban Informatics*.

Christin, D., Reinhardt, A., Kanhere, S. S. and Hollick, M. (2011) 'A survey on privacy in mobile participatory sensing applications', *Journal of Systems and Software*, 84(11), pp. 1928–1946. doi: 10.1016/j.jss.2011.06.073.

Christin, D. (2016) 'Privacy in mobile participatory sensing: Current trends and future challenges', *Journal of Systems and Software*, 116, pp. 57–68. doi: 10.1016/j.jss.2015.03.067.

Cicchino, J. B. (2017) 'Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates', *Accident Analysis and Prevention*, 99(2017), pp. 142–152. doi: 10.1016/j.aap.2016.11.009.

Cilliers, L. and Flowerday, S. (2015) 'The Relationship Between Privacy, Information Security and the Trustworthiness of a Crowdsourcing System in a Smart City', *Proceedings of the Ninth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2015)*.

Colonna, P., Intini, P., Berloco, N. and Ranieri, V. (2016) 'The influence of memory on driving behavior: How route familiarity is related to speed choice. An on-road study', *Safety Science*, 82, pp. 456–468. doi: 10.1016/j.ssci.2015.10.012.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M. and Hankey, J. (2016) 'Driver crash risk factors and prevalence evaluation using naturalistic driving data', *Proceedings of the National Academy of Sciences of the United States of America*, 113(10), pp. 2636–2641. doi: 10.1073/pnas.1513271113.

Eftekhari, H. R. and Ghatee, M. (2019) 'A similarity-based neuro-fuzzy modeling for driving behavior recognition applying fusion of smartphone sensors', *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 23(1), pp. 72–83. doi: 10.1080/15472450.2018.1506338.

Efthymiou, A., Barmpounakis, E. N., Efthymiou, D. and Vlahogianni, E. I. (2019) 'Transportation Mode Detection from Low-Power Smartphone Sensors Using Tree-Based Ensembles', *Journal of Big Data Analytics in Transportation*, 1(1), pp. 57–69. doi: https://doi.org/10.1007/s42421-019-00004-w.

El-Shawarby, I., Rakha, H., Inman, V. W. and Davis, G. W. (2007) 'Evaluation of driver deceleration behavior at signalized intersections', *Transportation Research Record*, (2018), pp. 29–35. doi: 10.3141/2018-05.

El-Shawarby, I., Rakha, H., Amer, A. and McGhee, C. (2011) 'Impact of Driver and Surrounding Traffic on Vehicle Deceleration Behavior at Onset of Yellow Indication', *Transportation research record*, 2248(1), pp. 10–20.

Elassad, Z. E. A., Mousannif, H., Al Moatassime, H. and Karkouch, A. (2020) 'The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review', *Engineering Applications of Artificial Intelligence*, 87(August 2019), p. 103312. doi: 10.1016/j.engappai.2019.103312.

Engelbrecht, J., Booysen, M. J., Van Rooyen, G. J. and Bruwer, F. J. (2016) 'Performance comparison of Dynamic Time Warping (DTW) and a Maximum Likelihood (ML) classifier in measuring driver behavior with smartphones', *Proceedings - 2015 IEEE Symposium Series on Computational Intelligence, SSCI 2015*, pp. 427–433. doi: 10.1109/SSCI.2015.70.

Etemad, M., Júnior, A. S. and Matwin, S. (2018) 'Predicting transportation modes of GPS trajectories using feature engineering and noise removal', in *Canadian Conference on Artificial Intelligence*. Springer, Cham, pp. 259–264. doi: 10.1007/978-3-319-89656-4_24.

Fafoutellis, P., Mantouka, E. G. and Vlahogianni, E. I. (2021) 'Eco-driving and its impacts on fuel efficiency: An overview of technologies and data-driven methods', *Sustainability (Switzerland)*, 13(1), pp. 1–17. doi: 10.3390/su13010226.

Fauvel, M., Chanussot, J. and Benediktsson, J. A. (2006) 'Decision fusion for the classification of urban remote sensing images', *IEEE Transactions on Geoscience and Remote Sensing*, 44(10), pp. 2828–2838. doi: 10.1109/TGRS.2006.876708.

Fazeen, M., Gozick, B., Dantu, R., Bhukhiya, M. and González, M. C. (2012) 'Safe Driving Using Mobile Phones', *IEEE Transactions on Intelligent Transportation Systems*, 13(3), pp. 1462–1468. doi: 10.1109/TITS.2012.2187640.

Fehr, E. and Fischbacher, U. (2003) 'The nature of human altruism', *Nature*, 425(6960), pp. 785–791. doi: 10.1038/nature02043.

Ferreira Júnior, J., Carvalho, E., Ferreira, B. V., De Souza, C., Suhara, Y., Pentland, A. and Pessin, G. (2017) 'Driver behavior profiling: An investigation with different smartphone sensors and machine learning', *PLoS ONE*, 12(4). doi: 10.1371/journal.pone.0174959.

Fujii, S. and Taniguchi, A. (2006) 'Determinants of the effectiveness of travel feedback programs-a review of communicative mobility management measures for changing travel behaviour in Japan', *Transport Policy*, 13(5), pp. 339–348. doi: 10.1016/j.tranpol.2005.12.007.

Ganti, R. K., Ye, F., Lei, H. and Watson, I. B. M. T. J. (2011) 'Mobile Crowdsensing : Current State and Future Challenges', *IEEE Communications Magazine*, (November), pp. 32–39.

Gärling, T. and Schuitema, G. (2007) 'Travel demand management targeting reduced private car use: Effectiveness, public acceptability and political feasibility', *Journal of Social Issues*, 63(1), pp. 139–153. doi: 10.1111/j.1540-4560.2007.00500.x.

Geroliminis, N. and Daganzo, C. F. (2007) 'Macroscopic Modeling of Traffic in Cities', in *Transportation Research Board*.

Geroliminis, N. and Daganzo, Carlos F. (2008) 'Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings', *Transportation Research Part B: Methodological*, 42(9), pp. 759–770. doi: 10.1016/j.trb.2008.02.002.

Geroliminis, N. and Daganzo, Carlos F (2008) 'Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings', *Transportation Research Part B: Methodological*, 42(9), pp. 759–770. doi: 10.11436/mssj.15.250.

Géron, A. (2017) *Hands-on Machine Learning with Scikit-Learn & TensorFlow: CONCEPTS, TOOLS, AND TECHNIQUES TO BUILD INTELLIGENT SYSTEMS*, *O'Reilly Media, Inc*.

Geyer, A., Kremslehner, D. and Muermann, A. (2019) 'Asymmetric Information in Automobile Insurance: Evidence From Driving Behavior', *Journal of Risk and Insurance*, (January). doi: 10.1111/jori.12279.

Gilman, E., Keskinarkaus, A., Tamminen, S., Pirttikangas, S., Roning, J. and Riekki, J. (2014) 'Personalised assistance for fuel-efficient driving', *Transportation Research Part C: Emerging Technologies*, 58, pp. 681–705. doi: 10.1016/j.trc.2015.02.007.

Gunduz, G., Yaman, C., Peker, A. U. and Acarman, T. (2018) 'Prediction of Risk Generated by Different Driving Patterns and Their Conflict Redistribution', *IEEE Transactions on Intelligent Vehicles*, 3(1), pp. 71–80. doi: 10.1109/TIV.2017.2788203.

Hall, D. L. and Llinas, J. (1997) 'An introduction to multisensor data fusion', *Proceedings of the IEEE*, 85(1), pp. 6–23. doi: 10.1109/5.554205.

Hamzeie, R., Savolainen, P. T. and Gates, T. J. (2017) 'Driver speed selection and crash risk : Insights from the naturalistic driving study', *Journal of Safety Research*, 63, pp. 187–194. doi: 10.1016/j.jsr.2017.10.007.

Handel, P., Skog, I., Wahlstrom, J., Bonawiede, F., Welch, R., Ohlsson, J. and Ohlsson, M. (2014) 'Insurance telematics: Opportunities and challenges with the smartphone solution', *IEEE Intelligent Transportation Systems Magazine*, 6(4), pp. 57–70. doi: 10.1109/MITS.2014.2343262.

Hasenjager, M., Heckmann, M. and Wersing, H. (2020) 'A Survey of Personalization for Advanced Driver Assistance Systems', *IEEE Transactions on Intelligent Vehicles*, 5(2), pp. 335–344. doi: 10.1109/TIV.2019.2955910.

Hasenjager, M. and Wersing, H. (2018) 'Personalization in advanced driver assistance systems and autonomous vehicles: A review', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2018-March, pp. 1–7. doi: 10.1109/ITSC.2017.8317803.

Haydari, A. and Yılmaz, Y. (2020) 'Deep Learning for Intelligent Transportation Systems: A Survey', *IEEE Transactions on Intelligent Transportation Systems*. doi: 10.1109/TITS.2019.2929020.

Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M. and Bhattacharya, J. (2007) 'Causality detection based on information-theoretic approaches in time series analysis', *Physics Reports*, 441(1), pp. 1–46. doi: 10.1016/j.physrep.2006.12.004.

Hoel, C. J., Wolff, K. and Laine, L. (2018) 'Automated Speed and Lane Change Decision Making using Deep Reinforcement Learning', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2018-Novem, pp. 2148–2155. doi: 10.1109/ITSC.2018.8569568.

Hong, J.-H., Margines, B. and Dey, A. K. (2014) 'A smartphone-based sensing platform to model aggressive driving behaviors', *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pp. 4047–4056. doi: 10.1145/2556288.2557321.

Hong, J. H., Margines, B. and Dey, A. K. (2014) 'A smartphone-based sensing platform to model aggressive driving behaviors', *Conference on Human Factors in Computing Systems - Proceedings*, (January), pp. 4047–4056. doi: 10.1145/2556288.2557321.

Hossain, M. (2012) 'Crowdsourcing: Activities, incentives and users' motivations to participate', *ICIMTR 2012 - 2012 International Conference on Innovation, Management and Technology Research*, (May 2012), pp. 501–506. doi: 10.1109/ICIMTR.2012.6236447.

Howe, J. (2006) 'The Rise of Crowdsourcing', *Wired Magazine*, 14(06), pp. 1–5. doi: 10.1086/599595.

Hsueh, P., Melville, P. and Sindhwani, V. (2009) 'Data quality from crowdsourcing: a study of annotation selection criteria', *Proceedings of the NAACL HLT 2009 Workshop on Active Learning for Natural Language Processing*, (June), pp. 27–35. doi: 10.1.1.157.5154.

Hu, X. L., Zhang, L. C. and Wang, Z. X. (2018) 'An adaptive smartphone anomaly detection model based on data mining', *Eurasip Journal on Wireless Communications and Networking*, 2018(1). doi: 10.1186/s13638-018-1158-6.

Huang, Y., Ng, E. C. Y., Zhou, J. L., Surawski, N. C., Chan, E. F. C. and Hong, G. (2018) 'Eco-driving technology for sustainable road transport: A review', *Renewable and Sustainable Energy Reviews*. doi: 10.1016/j.rser.2018.05.030.

Huang, Y., Santos, A. C. and Duhamel, C. (2020) 'Model and methods to address urban road network problems with disruptions', *International Transactions in Operational Research*, 27(6), pp. 2715–2739. doi: 10.1111/itor.12641.

Husnjak, S., Peraković, D., Forenbacher, I. and Mumdziev, M. (2015) 'Telematics system in usage based motor insurance', in *Procedia Engineering*, pp. 816–825. doi: 10.1016/j.proeng.2015.01.436.

Ignatov, A. (2018) 'Real-time human activity recognition from accelerometer data using Convolutional Neural Networks', *Applied Soft Computing Journal*, 62, pp. 915–922. doi: 10.1016/j.asoc.2017.09.027.

Jaritz, M., De Charette, R., Toromanoff, M., Perot, E. and Nashashibi, F. (2018) 'End-to-End Race Driving with Deep Reinforcement Learning', *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2070–2075. doi: 10.1109/ICRA.2018.8460934.

Jariyasunant, J., Carrel, A., Ekambaram, V., Gaker, D., Kote, T., Sengupta, R. and Walker, J. L. (2012) 'The Quantified Traveler: Using personal travel data to promote sustainable transport behavior', in *91st Annual Meeting Transportation Research Board*, pp. 1–15.

Johnson, D. A. and Trivedi, M. M. (2011) 'Driving style recognition using a smartphone as a sensor platform', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 1609–1615. doi: 10.1109/ITSC.2011.6083078.

Kanarachos, S., Christopoulos, S. R. G. and Chroneos, A. (2018) 'Smartphones as an integrated platform for monitoring driver behaviour: The role of sensor fusion and connectivity', *Transportation Research Part C: Emerging Technologies*, (March). doi: 10.1016/j.trc.2018.03.023.

Kanhere, S. S. (2013) 'Participatory Sensing: Crowdsourcing Data from Mobile Smartphones in Urban Spaces', *Mobile Data Management (MDM), 12th IEEE International Conference*, pp. 19–26.

Kanoje, S., Girase, S. and Mukhopadhyay, D. (2014) 'User Profiling Trends, Techniques and Applications', *International Journal of Advance Foundation and Research in Computer (IJAFRC)*, 1(1). Available at: http://arxiv.org/abs/1503.07474.

Kanungo, T., Member, S., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., Wu, A. Y. and Member, S. (2002) 'An Efficient k -Means Clustering Algorithm: Analysis and Implementation', 24(7), pp. 881–892.

Kaparias, I., Bell, M. G. . and Tomassini, M. (2011) 'Key Performance Indicators for traffic management and Intelligent Transport Systems', *ISIS*, 14(6). Available at: http://www.transport-research.info/Upload/Documents/201204/20120402_174223_62254_D 3.5 - Key Performance Indicators for traffic management and ITS.pdf.

Karlaftis, M. G. and Vlahogianni, E. I. (2011) 'Statistical methods versus neural networks in transportation research: Differences, similarities and some insights', *Transportation Research Part C: Emerging Technologies*, 19(3), pp. 387–399. doi: 10.1016/j.trc.2010.10.004.

Kaur, P. and Sobti, R. (2017) 'Current challenges in modelling advanced driver assistance systems: Future trends and advancements', *2017 2nd IEEE International Conference on*

Intelligent Transportation Engineering, ICITE 2017, pp. 236–240. doi: 10.1109/ICITE.2017.8056916.

Kaushik, M., Prasad, V., Krishna, K. M. and Ravindran, B. (2018) 'Overtaking Maneuvers in Simulated Highway Driving using Deep Reinforcement Learning', *IEEE Intelligent Vehicles Symposium, Proceedings*, 2018-June(Iv), pp. 1885–1890. doi: 10.1109/IVS.2018.8500718.

Kober, J., Bagnell, J. A. and Peters, J. (2013) 'Reinforcement learning in robotics: A survey', *International Journal of Robotics Research*, 32(11), pp. 1238–1274. doi: 10.1177/0278364913495721.

Kockelman, K. and Ma, J. (2018) 'Aggressive Driving and Speeding', in *Safe Mobility: Challenges, Methodology and Solutions*, pp. 37–55. doi: https://doi.org/10.1108/S2044-994120180000011003.

Koh, D. W. and Kang, H. B. (2015) 'Smartphone-based modeling and detection of aggressiveness reactions in senior drivers', in *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 12–17. doi: 10.1109/IVS.2015.7225655.

Kraft, P. (2012) 'Cluster analysis and cluster validation', (7), pp. 1–4.

Krajzewicz, Heinrich, M., Milano, M., Bellavista, P., Stützle, T., Härri, J. and Fellendorf, M. (2013) 'COLOMBO: Investigating the potential of V2X for traffic management purposes assuming low penetration rates', *ITS Europe*.

Krajzewicz, D., Hartinger, M., Hertkorn, G., Mieth, P., Rössel, C., Zimmer, J. and Wagner, P. (2005) 'Using the Road Traffic Simulation "SUMO" for Educational Purposes', in Hoogendoorn, S. P. et al. (eds) *Traffic and Granular Flow '03*. Springer, pp. 217–223. doi: 10.1007/3-540-28091-X.

Krajzewicz, D., Erdmann, J., Behrisch, M. and Bieker, L. (2012) 'Recent Development and Applications of {SUMO - Simulation of Urban MObility}', *International Journal On Advances in Systems and Measurements*, 5(3), pp. 128–138. doi: 10.1016/j.trc.2016.04.002.

Krajzewicz, D., Behrisch, M., Wagner, P., Luz, R. and Krumnow, M. (2015) 'Second generation of pollutant emission models for SUMO', *Modeling Mobility with Open Data*, pp. 203–221.

Lajunen, T. and Özkan, T. (2011) *Self-report instruments and methods*, *Handbook of traffic psychology*. Edited by B. E. Porter. London, UK: Elsevier.

Lavrenz, S. M., Vlahogianni, E. I., Gkritza, K. and Ke, Y. (2018) 'Time series modeling in traffic safety research', *Accident Analysis and Prevention*, 117(November 2017), pp. 368–380. doi: 10.1016/j.aap.2017.11.030.

Li, S., Li, P., Yao, Y., Han, X., Xu, Y. and Chen, L. (2020) 'Analysis of drivers' deceleration behavior based on naturalistic driving data', *Traffic Injury Prevention*, 21(1), pp. 42–47. doi: 10.1080/15389588.2019.1707194.

Liang, X., Wang, T., Yang, L. and Xing, E. (2018) 'CIRL: Controllable imitative reinforcement learning for vision-based self-driving', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11211 LNCS, pp.

604–620. doi: 10.1007/978-3-030-01234-2_36.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D. (2016) 'Continuous control with deep reinforcement learning', *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*.

Lin, K., Kansal, A., Lymberopoulos, D. and Zhao, F. (2010) 'Energy-accuracy trade-off for continuous mobile device location', *Proceedings of the 8th international conference on Mobile systems, applications, and services - MobiSys '10*, p. 285. doi: 10.1145/1814433.1814462.

Lin, N., Zong, C., Tomizuka, M., Song, P., Zhang, Z. and Li, G. (2014) 'An overview on study of identification of driver behavior characteristics for automotive control', *Mathematical Problems in Engineering*, (2). doi: 10.1155/2014/569109.

Liu, J., Cai, J., Lin, S. and Zhao, J. (2020) 'Analysis of Factors Affecting a Driver's Driving Speed Selection in Low Illumination', *Journal of Advanced Transportation*, 2020. doi: 10.1155/2020/2817801.

Liu, R., Zhao, X., Zhu, X. and Ma, J. (2019) 'Statistical characteristics of driver acceleration behaviour and its probability model', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*. doi: 10.1177/09544070211018039.

Liu, T., Yang, Y., Huang, G. Bin, Yeo, Y. K. and Lin, Z. (2016) 'Driver Distraction Detection Using Semi-Supervised Machine Learning', *IEEE Transactions on Intelligent Transportation Systems*, 17(4), pp. 1108–1120. doi: 10.1109/TITS.2015.2496157.

Liu, Z., Shen, Q. and Ma, J. (2017) 'A driving behavior model evaluation for UBI', *International Journal of Crowd Science*, 1(3), pp. 223–236. doi: 10.1108/ijcs-08-2017-0022.

Louta, M., Mpanti, K., Karetsos, G. and Lagkas, T. (2016) 'Mobile crowd sensing architectural frameworks: A comprehensive survey', in *IISA 2016 - 7th International Conference on Information, Intelligence, Systems and Applications*. doi: 10.1109/IISA.2016.7785385.

Magaña, V. C. and Organero, M. M. (2014) 'The impact of using gamificacion on the eco-driving learning', *Ambient Intelligence - Software and Applications, Advances in Intelligent Systems and Computing*, 291, pp. 45–52.

Magana, V. and Organero, M. (2011) 'Artemisa: Using an Android device as an Eco-Driving assistant', *Cyber Journals: Multidisciplinary Journals in Science and Technology: Journal of Selected Areas in Mechatronics (JMTC)*. Available at: http://www.cyberjournals.com/Papers/Jun2011/07.pdf.

Mairaj, A., Baba, A. I. and Javaid, A. Y. (2019) 'Application specific drone simulators: Recent advances and challenges', *Simulation Modelling Practice and Theory*, 94(October 2018), pp. 100–117. doi: 10.1016/j.simpat.2019.01.004.

Maldonado, S. and López, J. (2018) 'Dealing with high-dimensional class-imbalanced datasets: Embedded feature selection for SVM classification', *Applied Soft Computing Journal*, 67, pp. 94–105. doi: 10.1016/j.asoc.2018.02.051.

Mantouka, E., Barmpounakis, E., Vlahogianni, E. and Golias, J. (2020) 'Smartphone sensing for understanding driving behavior: Current practice and challenges', *International Journal of Transportation Science and Technology*. doi: 10.1016/j.ijtst.2020.07.001.

Mantouka, E. G., Barmpounakis, E. N. and Vlahogianni, E. I. (2018) 'Mobile Sensing and Machine Learning for Identifying Driving Safety Profiles', in *97th Annual Meeting Transportation Research Board*. Washington, D.C.

Mantouka, E. G., Barmpounakis, E. N. and Vlahogianni, E. I. (2019) 'Identifying driving safety profiles from smartphone data using unsupervised learning', *Safety Science*, 119(January), pp. 84–90. doi: 10.1016/j.ssci.2019.01.025.

Mashhadi, A. J. and Capra, L. (2011) 'Quality control for real-time ubiquitous crowdsourcing', *Proceedings of the 2nd international workshop on Ubiquitous crowdsouring*, pp. 5–8. doi: 10.1145/2030100.2030103.

Maulik, U. and Bandyopadhyay, S. (2002) 'Performance evaluation of some clustering algorithms and validity indices', *IEEE Transactions on pattern analysis and machine intelligence*, 24(12), pp. 1650–1654.

Meiring, G. A. M. and Myburgh, H. C. (2015) 'A review of intelligent driving style analysis systems and related artificial intelligence algorithms', *Sensors (Switzerland)*, 15(12), pp. 30653–30682. doi: 10.3390/s151229822.

Mensing, F., Bideaux, E., Trigui, R., Ribet, J. and Jeanneret, B. (2014) 'Eco-driving: An economic or ecologic driving style?', *Transportation Research Part C: Emerging Technologies*, 38, pp. 110–121. doi: 10.1016/j.trc.2013.10.013.

Meseguer, J. E., Toh, C. K., Calafate, C. T., Cano, J. C. and Manzoni, P. (2017) 'Drivingstyles: A mobile platform for driving styles and fuel consumption characterization', *Journal of Communications and Networks*, 19(2), pp. 162–168. doi: 10.1109/JCN.2017.000025.

Van Mierlo, J., Maggetto, G., Van De Burgwal, E. and Gense, R. (2004) 'Driving style and traffic measures - Influence on vehicle emissions and fuel consumption', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 218(1), pp. 43–50. doi: 10.1243/095440704322829155.

Min, K., Sim, G., Ahn, S., Sunwoo, M. and Jo, K. (2019) 'Vehicle deceleration prediction model to reflect individual driver characteristics by online parameter learning for autonomous regenerative braking of electric vehicles', *Sensors (Switzerland)*, 19(19). doi: 10.3390/s19194171.

Misra, A., Gooze, A., Watikins, K., Asad, M. and Le Dantec, C. A. (2014) 'Crowdsourcing and Its Application to Transportation Data Collection and Management', *Transportation Research Record: Journal of the Transportation Research Board*, 2(404), pp. 1–16. doi: 10.3141/2414-01.

Miyajima, B. C., Nishiwaki, Y., Ozawa, K., Wakita, T., Itou, K., Takeda, K. and Itakura, F. (2007) 'Driver Modeling Based on Driving Behavior and Its Evaluation in Driver Identification', in *Proceedings of the IEEE*, pp. 427–237.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M. (2013) 'Playing Atari with Deep Reinforcement Learning', *arXiv preprint arXiv:1312.5602.*, pp. 1–9. Available at: http://arxiv.org/abs/1312.5602.

Mohd, W. M. W., Beg, A. H., Herawan, T. and Rabbi, K. F. (2012) 'MaxD K-Means: A Clustering Algorithm for Auto-generation of Centroids and Distance of Data Points in Clusters', in *International Symposium on Intelligence Computation and Applications*. Berlin, Heidelberg: Springer, pp. 192–199. doi: 10.1007/978-3-642-34289-9.

Morales, M. (2020) *Grokking Deep Reinforcement Learning*. Manning Publications.

Möser, G. and Bamberg, S. (2008) 'The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence', *Journal of Environmental Psychology*, 28(1), pp. 10–26. doi: 10.1016/j.jenvp.2007.09.001.

Mukadam, M., Cosgun, A., Nakhaei, A. and Fujimura, K. (2017) 'Tactical Decision Making for Lane Changing with Deep Reinforcement Learning', *31st Conference on Neural Information Processing Systems (NIPS)*.

Musicant, O. and Lotan, T. (2016) 'Can novice drivers be motivated to use a smartphone based app that monitors their behavior?', *Transportation Research Part F: Traffic Psychology and Behaviour*, 42, pp. 544–557. doi: 10.1016/j.trf.2015.10.023.

Nandan, N., Pursche, A. and Zhe, X. (2014) 'Challenges in Crowdsourcing Real-Time Information for Public Transportation', *2014 IEEE 15th International Conference on Mobile Data Management*, 2, pp. 67–72. doi: 10.1109/MDM.2014.70.

Nikolic, M. and Bierlaire, M. (2017) 'Review of transportation mode detection approaches based on smartphone data', *17th Swiss Transport Research Conference*, (May).

Nowé, A. and Brys, T. (2016) 'A Gentle Introduction to Reinforcement Learning', in *International Conference on Scalable Uncertainty Management*. Springer, Cham, pp. 18–32. Available at: http://dx.doi.org/10.1007/978-3-642-23963-2.

Nuzzolo, A., Crisalli, U., Comi, A. and Rosati, L. (2014) 'An Advanced Traveller Advisory Tool Based on Individual Preferences', *Procedia - Social and Behavioral Sciences*, 160(Cit), pp. 539–547. doi: 10.1016/j.sbspro.2014.12.167.

Oltedal, S. and Rundmo, T. (2006) 'The effects of personality and gender on risky driving behaviour and accident involvement', *Safety Science*, 44(7), pp. 621–628. doi: 10.1016/j.ssci.2005.12.003.

Van Otterlo, M. and Wiering, M. (2012) 'Reinforcement learning and markov decision processes', *Adaptation, Learning, and Optimization*, 12, pp. 3–42. doi: 10.1007/978-3-642-27645-3_1.

Ouyang, Z., Niu, J. and Guizani, M. (2018) 'Improved Vehicle Steering Pattern Recognition by Using Selected Sensor Data', *IEEE Transactions on Mobile Computing*, 17(6), pp. 1383–1396. doi: 10.1109/TMC.2017.2762679.

Paefgen, J., Kehr, F., Zhai, Y. and Michahelles, F. (2012) 'Driving behavior analysis with smartphones: insights from a controlled field study', *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, (December), p. 36. doi: 10.1145/2406367.2406412.

Papadimitriou, E., Argyropoulou, A., Tselentis, D. I. and Yannis, G. (2019) 'Analysis of driver behaviour through smartphone data: The case of mobile phone use while driving', *Safety Science*, (January 2018), pp. 0–1. doi: 10.1016/j.ssci.2019.05.059.

Parasuraman, R. and Riley, V. (1997) 'Humans and Automation: Use, Misuse, Disuse, Abuse', *Human Factors*, 39(2). doi: 10.1518/001872097778543886.

Paúl, A., Fernández, J. A., Jiménez, F. and Sánchez, F. (2018) 'Chapter 7 - Cooperative Systems', in Jiménez, F. (ed.) *Intelligent Vehicles: Enabling Technologies and Future Developments*, pp. 227–274. doi: https://doi.org/10.1016/B978-0-12-812800-8.00007-2.

Pearson, S. (2013) 'Privacy, Security and Trust in Cloud Computing', *Privacy and Security for Cloud Computing*, pp. 3–42. doi: 10.1007/978-1-4471-4189-1.

Petridou, E. and Moustaki, M. (2000) 'Human factors in the causation of road traffic crashes', *European Journal of Epidemiology*, 16(9), pp. 819–826. doi: 10.1023/A:1007649804201.

Ponomarev, A. and Chernysheva, A. (2019) 'Adaptation and personalization in driver assistance systems', *Conference of Open Innovation Association, FRUCT*, 2019-April, pp. 335–344. doi: 10.23919/FRUCT.2019.8711961.

Predic, B. and Stojanovic, D. (2015) 'Enhancing driver situational awareness through crowd intelligence', *Expert Systems with Applications*, 42(11), pp. 4892–4909. doi: 10.1016/j.eswa.2015.02.013.

Richer, I. and Bergeron, J. (2012) 'Differentiating risky and aggressive driving: Further support of the internal validity of the Dula Dangerous Driving Index', *Accident Analysis and Prevention*, 45, pp. 620–627. doi: 10.1016/j.aap.2011.09.014.

Rittger, L., Schmidt, G., Maag, C. and Kiesel, A. (2015) 'Driving behaviour at traffic light intersections', *Cognition, Technology and Work*, 17(4), pp. 593–605. doi: 10.1007/s10111-015-0339-x.

Roy, A., Cruz, R. M. O., Sabourin, R. and Cavalcanti, G. D. C. (2018) 'A study on combining dynamic selection and data preprocessing for imbalance learning', *Neurocomputing*, 286, pp. 179–192. doi: 10.1016/j.neucom.2018.01.060.

Sadia, R., Bekhor, S. and Polus, A. (2018) 'Speed variation for different drivers, situations, and road geometry: Simulator and survey analysis', *Journal of Transportation Safety and Security*, 10(1–2), pp. 25–44. doi: 10.1080/19439962.2016.1275071.

Sagberg, F., Selpi, P. G. F. B. and Engstrom, J. (2015) 'A Review of Research on Driving Styles and Road Safety', *Human Factors The Journal of the Human Factors and Ergonomics Society*, 57(7), pp. 1248–1275. doi: 10.1177/0018720815591313.

Saiprasert, C. and Pattara-Atikom, W. (2013) 'Smartphone enabled dangerous driving report system', *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 1231–1237. doi: 10.1109/HICSS.2013.484.

Saiprasert, C., Pholprasit, T. and Pattara-Atikom, W. (2013) 'Detecting Driving Events using Smartphone', *20th ITS World Congress*, (October), pp. 1–12. Available at: http://trid.trb.org.ezproxy.library.wisc.edu/view/2013/C/1323676%255Cnhttps://drive.google.com/open?id=0B1-iNPy2dfV0dElXUWdUUEJZdzg&authuser=0.

Saiprasert, C., Pholprasit, T. and Thajchayapong, S. (2017) 'Detection of Driving Events using Sensory Data on Smartphone', *International Journal of Intelligent Transportation Systems Research*, 15(1), pp. 17–28. doi: 10.1007/s13177-015-0116-5.

Sallab, A. El, Abdou, M., Perot, E. and Yogamani, S. (2017) 'Deep reinforcement learning framework for autonomous driving', *arXiv*, pp. 70–76.

Sarwar, A. and Khan, M. (2013) 'A Review of Trust Aspects in Cloud Computing Security', *International Journal of Cloud Computing …*, 2(2), pp. 116–122. doi: http://dx.doi.org/10.11591/closer.v2i2.1840.

Schmitt, M. and Zhu, X. X. (2016) 'Data Fusion and Remote Sensing: An ever-growing relationship', *IEEE Geoscience and Remote Sensing Magazine*, 4(4), pp. 6–23. doi: 10.1109/MGRS.2016.2561021.

Seger, C. (2018) 'An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing', *Degree Project Technology*, p. 41. Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-237426%0Ahttp://www.diva-portal.org/smash/get/diva2:1259073/FULLTEXT01.pdf.

Semenescu, A., Gavreliuc, A. and Sârbescu, P. (2020) '30 Years of soft interventions to reduce car use – A systematic review and meta-analysis', *Transportation Research Part D: Transport and Environment*, 85(June), p. 102397. doi: 10.1016/j.trd.2020.102397.

Shi, Q., Zhang, Y. L., Li, L., Yang, X., Li, M. and Zhou, J. (2020) 'SAFE: Scalable automatic feature engineering framework for industrial tasks', *Proceedings - International Conference on Data Engineering*, 2020-April, pp. 1645–1656. doi: 10.1109/ICDE48307.2020.00146.

Shukla, A. K., Sharma, R. and Muhuri, P. K. (2018) 'A review of the scopes and challenges of the modern real-time operating systems', *International Journal of Embedded and Real-Time Communication Systems*, 9(1), pp. 66–82. doi: 10.4018/IJERTCS.2018010104.

Silver, D., Lever, G., Heess, N., Degris, T. and Wierstra, Daan, Riedmiller, M. (2014) 'Deterministic policy gradient algorithms', *ICML*.

Simons-Morton, B. G., Ouimet, M. C., Zhang, Z., Klauer, S. E., Lee, S. E., Wang, J., Chen, R., Albert, P. and Dingus, T. A. (2011) 'The effect of passengers and risk-taking friends on risky driving and crashes/near crashes among novice teenagers', *Journal of Adolescent Health*, 49(6), pp. 587–593. doi: 10.1016/j.jadohealth.2011.02.009.

Singh, G., Bansal, D. and Sofat, S. (2017) 'A smartphone based technique to monitor driving behavior using DTW and crowdsensing', *Pervasive and Mobile Computing*, 40, pp. 56–70. doi: 10.1016/j.pmcj.2017.06.003.

Smith, J. G., Ponnuru, S. K. and Patil, M. (2016) 'Detection of aggressive driving behavior and fault behavior using pattern matching', in *2016 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2016*. Jaipur, India: IEEE, pp. 207–211. doi: 10.1109/ICACCI.2016.7732048.

Staubach, M., Schebitz, N., Köster, F. and Kuck, D. (2014) 'Evaluation of an eco-driving support system', *Transportation Research Part F: Traffic Psychology and Behaviour*, 27(PA), pp. 11–21. doi: 10.1016/j.trf.2014.09.006.

Stavrakaki, A. M., Tselentis, D. I., Barmpounakis, E., Vlahogianni, E. I. and Yannis, G. (2020) 'Estimating the necessary amount of driving data for assessing driving behavior', *Sensors (Switzerland)*, 20(9). doi: 10.3390/s20092600.

Sutton, R. S. and Barto, A. G. (2018) *Reinforcement Learning. An introduction*. MIT Press.

Tang, X., Chen, Y., Li, X., Liu, J. and Ying, Z. (2019) 'A reinforcement learning approach to personalized learning recommendation systems', *British Journal of Mathematical and Statistical Psychology*, 72, pp. 108–135. doi: 10.1111/bmsp.12144.

Tasca, L. (2000) *A Review of the Literature on Aggressive Driving Research*. Ontario, Canada: Ontario Advisory Group on Safe Driving Secretariat Road User Safety Branch, Ontario Ministry of Transportation. doi: 10.1002/aehe.3640230702.

Tornros, J. E. B. and Bolling, A. K. (2005) 'Mobile phone use — Effects of handheld and handsfree phones on driving performance', *Accident Analysis & Prevention*, 37, pp. 902–909. doi: 10.1016/j.aap.2005.04.007.

Tselentis, D. I. (2018) *Benchmarking Driving Efficiency using Data Science Techniques applied on Large-Scale Smartphone Data*.

Tselentis, D. I., Yannis, G. and Vlahogianni, E. I. (2016) 'Innovative Insurance Schemes: Pay as/how You Drive', *Transportation Research Procedia*, 14, pp. 362–371. doi: 10.1016/j.trpro.2016.05.088.

Tselentis, D. I., Yannis, G. and Vlahogianni, E. I. (2017) 'Innovative motor insurance schemes: A review of current practices and emerging challenges', *Accident Analysis and Prevention*, 98, pp. 139–148. doi: 10.1016/j.aap.2016.10.006.

Tuyen, L. P. and Chung, T. C. (2017) 'Controlling bicycle using deep deterministic policy gradient algorithm', *14th International Conference on Ubiquitous Robots and Ambient Intelligence, URAI 2017*, pp. 413–417. doi: 10.1109/URAI.2017.7992765.

United States Department of Education (2017) *Reimagining the role of technology in education: 2017 National Education Technology Plan update*. Washington, DC. Available at: https://tech.ed.gov/files/2017/01/NETP17.pdf.

Vaiana, R., Iuele, T., Astarita, V., Caruso, M. V., Tassitani, A., Zaffino, C. and Giofrè, V. P. (2014) 'Driving behavior and traffic safety: An acceleration-based safety evaluation procedure for smartphones', *Modern Applied Science*, 8(1), pp. 88–96. doi: 10.5539/mas.v8n1p88.

Varmuza, K. and Filzmoser, P. (2016) *Introduction to Multivariate Statistical Analysis in Chemometrics*. CRC press.

Vlachogiannis, D. M., Vlahogianni, E. I. and Golias, J. (2020) 'A reinforcement learning model for personalized driving policies identification', *International Journal of Transportation Science and Technology*, 9(4), pp. 299–308. doi: 10.1016/j.ijtst.2020.03.002.

Vlahogianni, E. I. and Barmpounakis, E. N. (2017a) 'Driving analytics using smartphones: Algorithms, comparisons and challenges', *Transportation Research Part C: Emerging Technologies*, 79, pp. 196–206. doi: 10.1016/j.trc.2017.03.014.

Vlahogianni, E. I. and Barmpounakis, E. N. (2017b) 'Gamification and Sustainable Mobility: Challenges and Opportunities in a Changing Transportation Landscape', in *Low Carbon Mobility for Future Cities: Principles and Applications*.

Wahlström, J., Skog, I. and Händel, P. (2015) 'Detection of Dangerous Cornering in GNSS-Data-Driven Insurance Telematics', *IEEE Transactions on Intelligent Transportation Systems*, 16(6), pp. 3073–3083. doi: 10.1109/TITS.2015.2431293.

Wahlström, J., Skog, I. and Händel, P. (2017) 'Smartphone-Based Vehicle Telematics: A Ten-Year Anniversary', *IEEE Transactions on Intelligent Transportation Systems*, 18(10), pp. 2802–2825. doi: 10.1109/TITS.2017.2680468.

Wang, H., Bah, M. J. and Hammad, M. (2019) 'Progress in Outlier Detection Techniques: A Survey', *IEEE Access*, 7, pp. 107964–108000. doi: 10.1109/ACCESS.2019.2932769.

Wang, J., Wang, Y., Zhang, D. and Helal, S. (2018) 'Energy Saving Techniques in Mobile Crowd Sensing: Current State and Future Opportunities', *IEEE Communications Magazine*, 56(5), pp. 164–169. doi: 10.1109/MCOM.2018.1700644.

Wang, P., Chan, C. Y. and De La Fortelle, A. (2018) 'A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers', *IEEE Intelligent Vehicles Symposium, Proceedings*, 2018-June(Iv), pp. 1379–1384. doi: 10.1109/IVS.2018.8500556.

Wang, P., Li, H. and Chan, C. Y. (2019) 'Continuous control for automated lane change behavior based on deep deterministic policy gradient algorithm', in *IEEE Intelligent Vehicles Symposium*. IEEE, pp. 1454–1460.

Wang, S., Jia, D. and Weng, X. (2018) *Deep Reinforcement Learning for Autonomous Driving*. doi: 10.1007/978-3-030-47358-7_7.

Wang, Y., Jia, X., Jin, Q. and Ma, J. (2017) 'Mobile crowdsourcing: framework, challenges, and solutions', *Concurrency Computation Practice and Experience*, 29(3). doi: 10.1002/cpe.

Wells, S., Kotkanen, H., Schlafli, M. and Gabrielli, S. (2014) 'Towards an Applied Gamification Model for Tracking, Managing, & Encouraging Sustainable Travel Behaviours', *ICST Trans.*

Ambient Systems, 1(4), pp. 1–15. doi: 10.4108/amsys.1.4.e2.

Westin, A. F. (1967) *Privacy and Freedom*. New York: Atheneum.

Wickham, H. (2014) 'Tidy Data', *Journal of Statistical Software*, 59(10), pp. 1–23. Available at: http://www.jstatsoft.org/.

Xintong, G., Hongzhi, W., Song, Y. and Hong, G. (2014) 'Brief survey of crowdsourcing for data mining', *Expert Systems with Applications*, 41(17), pp. 7987–7994. doi: 10.1016/j.eswa.2014.06.044.

Yan, L., Jiang, L., Xia, Y. and Fu, M. (2016) 'State estimation and data fusion for multirate sensor networks', *International Journal of Adaptive Control and Signal Processing*, 30, pp. 3–15. doi: 10.1002/acs.

Yang, D., Xue, G., Fang, X. and Tang, J. (2016) 'Incentive Mechanisms for Crowdsensing: Crowdsourcing with Smartphones', *IEEE/ACM Transactions on Networking*, 24(3), pp. 1732–1744. doi: 10.1109/TNET.2015.2421897.

Yang, K., Zhang, K., Ren, J. and Shen, X. (2015) 'Security and privacy in mobile crowdsourcing networks: challenges and opportunities', *IEEE Communications Magazine*, 53(8), pp. 75–81. doi: 10.1109/MCOM.2015.7180511.

Yen, B. T. H., Mulley, C. and Burke, M. (2019) 'Gamification in transport interventions: Another way to improve travel behavioural change', *Cities*, 85(July 2018), pp. 140–149. doi: 10.1016/j.cities.2018.09.002.

Yi, D., Su, J., Liu, C., Quddus, M. and Chen, W. H. (2019) 'A machine learning based personalized system for driving state recognition', *Transportation Research Part C: Emerging Technologies*, 105(July 2018), pp. 241–261. doi: 10.1016/j.trc.2019.05.042.

You, C., Lu, J., Filev, Di. and Tsiotras, P. (2018) 'Highway Traffic Modeling and Decision Making for Autonomous Vehicle Using Reinforcement Learning', *IEEE Intelligent Vehicles Symposium, Proceedings*, 2018-June(Iv), pp. 1227–1232. doi: 10.1109/IVS.2018.8500675.

Yu, J., Chen, Z., Zhu, Y., Jennifer Chen, Y., Kong, L. and Li, M. (2017) 'Fine-Grained Abnormal Driving Behaviors Detection and Identification with Smartphones', *IEEE Transactions on Mobile Computing*, 16(8), pp. 2198–2212. doi: 10.1109/TMC.2016.2618873.

Yuan, Y. and Raubal, M. (2012) 'Extracting dynamic urban mobility patterns from mobile phone data', *Geographic information science*, pp. 354–367. doi: 10.1007/978-3-642-33024-7_26.

Zahid, M., Chen, Y., Khan, S., Jamal, A., Ijaz, M. and Ahmed, T. (2020) 'Predicting risky and aggressive driving behavior among taxi drivers: Do spatio-temporal attributes matter?', *International Journal of Environmental Research and Public Health*, 17(11), pp. 1–21. doi: 10.3390/ijerph17113937.

Zhang, H., Sun, J. and Tian, Y. (2020) 'The impact of socio-demographic characteristics and driving behaviors on fuel efficiency', *Transportation Research Part D: Transport and Environment*, 88(September), p. 102565. doi: 10.1016/j.trd.2020.102565.

Zhang, Y., Sun, P., Yin, Y., Lin, L. and Wang, X. (2018) 'Human-like Autonomous Vehicle Speed Control by Deep Reinforcement Learning with Double Q-Learning', *IEEE Intelligent Vehicles Symposium, Proceedings*, 2018-June(Iv), pp. 1251–1256. doi: 10.1109/IVS.2018.8500630.

Zhao, X., Xu, W., Ma, J., Li, H. and Chen, Y. (2019) 'An analysis of the relationship between driver characteristics and driving safety using structural equation models', *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, pp. 529–545. doi: 10.1016/j.trf.2019.02.004.

Zheng, A. and Casari, A. (2018) *Feature engineering for machine learning: principles and techniques for data scientists*. O'Reilly Media, Inc.

Ziakopoulos, A., Tselentis, D., Kontaxi, A. and Yannis, G. (2020) 'A critical overview of driver recording tools', *Journal of Safety Research*, 72, pp. 203–212. doi: 10.1016/j.jsr.2019.12.021.

Zolali, M., Mirbaha, B., Layegh, M. and Behnood, H. R. (2021) 'A Behavioral Model of Drivers' Mean Speed Influenced by Weather Conditions, Road Geometry, and Driver Characteristics Using a Driving Simulator Study', *Advances in Civil Engineering*, 2021.

Appendix I

List of publications produced within the framework of this dissertation

❖ Scientific journals:

Mantouka E. & Vlahogianni E., (2021). Deep Reinforcement Learning for Personalized Driving Recommendations: Modeling and Impact Assessment, Transportation Research Part C: Emerging Technologies (under review)

Mantouka E., Barmpounakis E., Vlahogianni E. & Golias J. (2020). Smartphone sensing for understanding driving behavior: Current practice and challenges. International Journal of Transportation Science and Technology, vol. 10(3), 266-282, https://doi.org/10.1016/j.ijtst.2020.07.001. [Citations: 6]

Adamidis F. K., Mantouka E. G., & Vlahogianni E. I. (2020), Effects of controlling aggressive driving behavior on network-wide traffic flow and emissions. International Journal of Transportation Science and Technology, vol. 9(3), pp. 263-276, https://doi.org/10.1016/j.ijtst.2020.05.003. [Citations: 9]

Mantouka E., Barmpounakis E. & Vlahogianni E. (2019). Identifying driving safety profiles from smartphone data using unsupervised learning, Safety Science, vol. 119, pp. 84-90, https://doi.org/10.1016/j.ssci.2019.01.025 [Citations: 31]

❖ Scientific conferences:

Mantouka E. & Vlahogianni E., (2021). Deep Reinforcement Learning for Controlling Driving Behavior and Its Impact on Traffic and Road Safety, In the 101th Annual Meeting Transportation Research Board (TRB), Washington D.C., US (accepted for presentation)

Mantouka E.G., Vlahogianni E. & Golias I. (2019). What is your driving identity? Some empirical findings using large-scale smartphone sensors' data, In the 9th International Congress on Transportation Research (ICTR), Αθήνα, Ελλάδα

Mantouka E. G., Barmpounakis E. N. & Vlahogianni E. I. (2018). Mobile Sensing and Machine Learning for Identifying Driving Safety Profiles, In the 97th Annual Meeting Transportation Research Board (TRB), Washington D.C., US [Citations: 11]

Mantouka E. G., Barmpounakis E. N. & Vlahogianni E. I. (2017). Methodological Issues and Concerns from Using Smartphones to Collect Driving and Mobility Data, In the 8th International Congress on Transportation Research (ICTR), Θεσσαλονίκη, Ελλάδα

List of publications in other research thematic areas

❖ Scientific journals:

1. Fafoutellis P., Mantouka E., Vlahogianni E.I., Oprea G.M. (2021). Acceptability modeling of autonomous mobility on-demand services with on-board ride sharing using interpretable Machine Learning. International Journal of Transportation Science and Technology, In press, https://doi.org/10.1016/j.ijtst.2021.10.003
2. Mantouka E., Fafoutellis P., Vlahogianni E., Oprea G.M. (2021). Understanding User Perception and Feelings for Autonomous Mobility on Demand in the COVID-19 Pandemic Era. Transportation Research Interdisciplinary Perspectives (under review)
3. Mantouka E., Fafoutellis P., Vlahogianni E. (2021). Acceptance of a Pay-How-You-Drive Pricing Scheme for City Traffic: The Case of Athens. Transportation Research Part A: Policy and Practice (under review)
4. Mourtakos V., Fafoutellis P., Mantouka E., Vlahogianni E. (2021). Reconstructing Mobility from Smartphone Data: Empirical Evidence of the Effects of COVID-19 Pandemic Crisis on Working and Leisure, Transport Policy (under review)
5. Fafoutellis P., Mantouka E., Vlahogianni E.I. (2021), Deep survival analysis of searching for on-street parking in urban areas. Transportation Research Part C: Emerging Technologies, 128, 103173, https://doi.org/10.1016/j.trc.2021.103173
6. Fafoutellis P., Mantouka E., Vlahogianni E.I. (2020). Eco-driving and its impacts on fuel efficiency: An overview of technologies and data-driven methods, Sustainability, 13 (1), 226.
7. Mantouka E. G., Papacharalampous A., Heydenrijk-Ottens L., Shelat S., Degeler V. & van Lint H. (2019). Understanding Travel Behavior Through Travel Happiness, Transportation Research Record, vol. 2673 (4), pp. 889-897
8. Mantouka E. G., Barmpounakis E. N., Milioti C. P. & Vlahogianni E. I. (2018), Gamification in mobile applications: The case of airports, Journal of Intelligent Transportation Systems, vol. 23 (5), pp. 417-426
9. Christovasili K., Mantouka E., Vlahogianni E. (2021) A User Acceptance Survey of Pay-How-You-Drive Urban Pricing Schemes. In: Nathanail E.G., Adamos G., Karakikes I. (eds) Advances in Mobility-as-a-Service Systems. CSUM 2020. Advances in Intelligent Systems and Computing, vol 1278. Springer, Cham. https://doi.org/10.1007/978-3-030-61075-3_57
10. Fafoutellis P. & Mantouka E. G. (2019), Major Limitations and Concerns Regarding the Integration of Autonomous Vehicles in Urban Transportation Systems, Data Analytics: Paving the Way to Sustainable Urban Mobility. CSUM 2018. Advances in Intelligent Systems and Computing, vol 879, Nathanail E., Karakikes I. (eds). Springer, Cham

❖ Scientific conferences:

11. Mourtakos V., Fafoutellis P., Mantouka E., Vlahogianni E. (2021). Extracting Activity Patterns and Trip Chains from Smartphone Data: A Case Study on the effects of COVID-

19 on Greek Mobility, In the 101th Annual Meeting Transportation Research Board (TRB), Washington D.C., US (accepted for presentation)

12. Konstantinou C., Mourtakos V., Fafoutellis P., Mantouka E., Vlahogianni E. (2021). Mining Human Mobility Patterns, Activity Chains and Duration using Smartphone Sensing, In the 10th International Congress on Transportation Research (ICTR), Ρόδος, Ελλάδα
13. Haxhi A., Perakis H., Gikas V., Fafoutellis P., Mantouka E., Vlahogianni E., Pnevmatikou A., Fortsakis P. (2021). Towards Smartphone-based enhanced GNSS positioning for Eco Driving ITS services, In the 10th International Congress on Transportation Research (ICTR), Ρόδος, Ελλάδα
14. Fafoutellis P., Mantouka E., Vlahogianni E., Haxhi A., Perakis H., Gikas V., Pnevmatikou A., Kostoulas G., Frantzola E. (2021). What is the Impact of Driving Behavior on Fuel Efficiency? Theoretical Aspects and Empirical Evidence, In the 10th International Congress on Transportation Research (ICTR), Ρόδος, Ελλάδα
15. Christovasili K., Mantouka E., Vlahogianni E. (2020). A User Acceptance Survey of Pay-How-You-Drive Urban Pricing Schemes, In the 5th Conference On Sustainable Urban Mobility (CSUM), Virtual Event
16. Mantouka E., Fafoutellis P., Vlahogianni E. (2020). Data Science Models for Analyzing Cruising for On-Street Parking, In the 99th Annual Meeting Transportation Research Board (TRB), Washington D.C., US
17. Mantouka E., Lagou G., Barmpounakis E. & Vlahogianni E. (2019), Mapping Risky Driving Behavior in Urban Road Networks, In the 9th International Congress on Transportation Research (ICTR), Αθήνα, Ελλάδα
18. Mantouka E. G., Drivakou A., Moschovou T. & Vlahogianni E.I. (2019). European Citizens' Mindset versus European Commission's Targets for Alternative Fuels, In the 9th International Congress on Transportation Research (ICTR), Αθήνα, Ελλάδα
19. Mantouka E. G., Koliou P., Papacharalampous A., Vlahogianni E.I. & Deloukas A. (2019). Revisiting travel mode and time of departure choices in EU regions: Differences, similarities and some insights, In the 9th International Congress on Transportation Research (ICTR), Αθήνα, Ελλάδα
20. Mantouka E. G., Orfanou F, Margreiter M. & Vlahogianni E.I. (2019). Smart Parking Assistance Services and User Acceptance: A European Model, In the 5th International Conference on Vehicle Technology and Intelligent Transport Systems
21. Adamidis F., Mantouka E. G., Barmpounakis E. N. & Vlahogianni E. I. (2018). Impacts of Eco Driving on Traffic Flow and Emissions in Large Scale Urban Networks, In the 98th Annual Meeting Transportation Research Board (TRB), Washington D.C., US
22. Toulouki M., Mantouka E. G., Vlahogianni E. I., Gkritza K. & Kepaptsoglou K. (2018). Perceived Impacts of c- ITS on the Economy, Quality of Life and Transportation System Performance: The Case of Greece, In the 98th Annual Meeting Transportation Research Board (TRB), Washington D.C., US

23. Mantouka E. G., Papacharalampous A., Heydenrijk-Ottens L., Shelat S., Degeler V. & van Lint H. (2018). Understanding Travel Behavior Through Travel Happiness, In the 98th Annual Meeting Transportation Research Board (TRB), Washington D.C., US
24. Mantouka E. G., Vlahogianni E. I., Papacharalampous A., Heydenrijk-Ottens L., Shelat S., Chamoso Santos P., Guisado-Gamez J., Mitsakis E., Degeler V., & Larriba-Pey J. L. (2018). Urban Travel Behaviour: A Cross-country Comparison, In the 7th Symposium of the European Association for Research in Transportation conference (hEART), 5-7 Σεπτεμβρίου, Αθήνα, Ελλάδα
25. Di Pasquale G., Degeler V., Papacharalampous A., Mantouka E. G., Larriba-Pey J.L. & Vlahogianni E.I. (2018). Multimodal Travel Companion Enabled by Artificial Intelligence, In the 25th ITS World Congress, 17-21 Σεπτεμβρίου, Κοπεγχάγη, Δανία
26. Mantouka E. G., Barmpounakis E. N. & Vlahogianni E. I. (2016). Mobile Apps for Airports: A User Acceptance Analysis, European Conference on Mobility Management (ECOMM), Αθήνα, Ελλάδα