



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ & ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

**“Μεθοδολογία Εκτίμησης του Απαιτούμενου Όγκου Πληθοποριστικών
Δεδομένων Οδήγησης για την Αξιολόγηση της Οδηγικής
Συμπεριφοράς”**

Διπλωματική Εργασία



Άννα – Μαρία Σταυρακάκη

Επιβλέπουσα: Ελένη Ι. Βλαχογιάννη

Αθήνα, Ιούλιος 2018

Ευχαριστίες

Ολοκληρώνοντας τη διπλωματική μου εργασία, θα ήθελα να ευχαριστήσω όσους με βοήθησαν και με στήριξαν σε αυτή τη πενταετή προσπάθειά μου. Αρχικά, θα ήθελα να ευχαριστήσω την επιβλέπουσα Καθηγήτριά μου, κυρία Ελένη Βλαχογιάννη, για την εμπιστοσύνη που μου έδειξε με την ανάθεση της παρούσας Διπλωματικής Εργασίας, για την υποστήριξή της καθώς και για την εκπληκτική συνεργασία που είχαμε σε όλα τα στάδια εκπόνησής της. Επιπλέον, οφείλω να την ευχαριστήσω για τη σωστή καθοδήγηση αλλά και για τις γενικότερες γνώσεις που μου μετέδωσε, συμβάλλοντας καθοριστικά στην αγάπη που έχω αναπτύξει για το αντικείμενο του Συγκοινωνιολόγου Μηχανικού.

Επίσης, οφείλω ένα μεγάλο ευχαριστώ στον κ. Δημήτρη Τσελέντη, Υποψήφιο Διδάκτορα ΕΜΠ, ο οποίος ήταν έτοιμος να απαντήσει στις ερωτήσεις μου κάθε φορά που χρειάστηκε, αλλά και για τις παρατηρήσεις και τις συμβουλές του κατά τη διάρκεια εκπόνησης της παρούσας Διπλωματικής Εργασίας. Χωρίς τη βοήθειά του δε θα τα είχα καταφέρει ως εδώ. Ιδιαίτερες ευχαριστίες θα ήθελα να απευθύνω και στους λοιπούς υποψήφιους διδάκτορες του Εργαστηρίου Κυκλοφοριακής Τεχνικής. Εκεί βρήκα έναν πολύτιμο σύμμαχο, τη στιγμή που τον χρειάστηκα.

Ιδιαίτερες ευχαριστίες ανήκουν και στην εταιρεία OSeven, η οποία παραχώρησε τα πολύ χρήσιμα στοιχεία μέτρησης της οδηγικής συμπεριφοράς που χρησιμοποιήθηκαν στην παρούσα Διπλωματική Εργασία.

Τέλος, θα ήθελα να εκφράσω τις θερμές ευχαριστίες μου προς τους φίλους και τους συμφοιτητές μου, για τις πολύ όμορφες αλλά και δύσκολες στιγμές που περάσαμε μαζί όλα αυτά τα χρόνια. Μα, πιο πολύ θα ήθελα να ευχαριστήσω την οικογένειά μου για την αγάπη και την αμέριστη συμπαράστασή της, ηθική και υλική, από τα πρώτα μου βήματα, μέχρι τώρα και για πάντα.

Αθήνα, Ιούλιος 2018
Άννα-Μαρία Σταυρακάκη

**“ Μεθοδολογία Εκτίμησης του Απαιτούμενου Όγκου Πληθοποριστικών
Δεδομένων Οδήγησης για την Αξιολόγηση της Οδηγικής
Συμπεριφοράς”**

Άννα-Μαρία Σταυρακάκη

Επιβλέπουσα: Ελένη Ι. Βλαχογιάννη, Επίκουρη καθηγήτρια ΕΜΠ

Σύνοψη

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μεθοδολογίας για την εκτίμηση του εύρους δεδομένων οδήγησης που απαιτείται να συλλέγονται για κάθε οδηγό για τον ακριβή προσδιορισμό της οδηγικής του συμπεριφοράς. Συνολικά συλλέχθηκαν δεδομένα για 21.610 διαδρομές από 68 οδηγούς. Για τον προσδιορισμό του εύρους των απαραίτητων δεδομένων για τη σύγκλιση της οδηγικής συμπεριφοράς, χρησιμοποιήθηκαν σωρευτικά αθροίσματα των απότομων συμβάντων επιτάχυνσης και επιβράδυνσης, αλλά και του χρόνου χρήσης κινητού τηλεφώνου και υπέρβασης του ορίου ταχύτητας, καθώς και η μεταβλητότητα αυτών, σε συνδυασμό με απλά στατιστικά όρια και κινητούς μέσους όρους. Τα αποτελέσματα έδειξαν ότι δεν υπάρχει ένα διακριτό χρονικό σημείο, κοινό για όλους τους οδηγούς ή/και όλα τα χαρακτηριστικά, από το οποίο και μετά η οδηγική συμπεριφορά να σταθεροποιείται. Συμπεραίνεται ότι για τον καθορισμό του χρονικού διαστήματος παρατήρησης ενός οδηγού, θα πρέπει να έχει προηγηθεί μια ανάλυση του προφίλ επιθετικότητας αλλά και μεταβλητότητας της συμπεριφοράς του από ταξίδι σε ταξίδι.

Λέξεις κλειδιά: οδηγική συμπεριφορά, έξυπνα κινητά τηλέφωνα, επιτάχυνση, επιβράδυνση, απότομα συμβάντα, σύγκλιση, μεταβλητότητα, σωρευτικά αθροίσματα, κινητός μέσος όρος

“Methodology for Assessing the Required Volume of Driving Data for Assessing Driving Behavior”

Anna-Maria Stavrakaki

Supervisor: Eleni I. Vlahogianni, Assistant Professor NTUA

Abstract

The aim of this diploma thesis is to develop a methodology for estimating the amount of driving data to be collected for each driver to accurately determine his driving behavior. A total of 21,610 routes were collected from 68 drivers. To determine the necessary data for the convergence of driving behavior, a methodological approach based on cumulative sums of driving behavior metrics (harsh acceleration/braking events, mobile usage and speeding), as well as their volatility, combined with simple statistical measures and moving averages, is developed. The results showed that there is not one distinct number of trips, common for all drivers and / or all the metrics, after which the driving behavior stabilizes. Therefore, in order to determine exactly how much driving data should be recorded from each driver, an analysis of the driver's aggressiveness and his behavior volatility should be conducted.

Keywords: driving behavior, smartphones, acceleration, braking, harsh events, convergence, volatility, cumulative sum, moving average

Περίληψη

Οι ανθρώπινοι παράγοντες θεωρούνται μια από τις κύριες αιτίες θανάτων από τροχαία ατυχήματα και τραυματισμούς κάθε χρόνο, και ως εκ τούτου είναι εξαιρετικά σημαντικό να μελετηθεί πώς αυτοί οι παράγοντες μπορούν να επηρεάσουν τον κυκλοφοριακό κίνδυνο. Η ταχεία τεχνολογική πρόοδος, ειδικά στην Τηλεματική, καθώς και η αύξηση της διείσδυσης και της χρήσης των τεχνολογιών πληροφόρησης από τους οδηγούς (π.χ. Smartphones), μπορούν να συμβάλλουν στην βαθύτερη κατανόηση και πρόληψη των παραγόντων εκείνων που οδηγούν στα οδικά ατυχήματα, μέσω των ευκαιριών που προσφέρουν για την ακριβή παρακολούθηση, καταγραφή, ανάλυση και αξιολόγηση της συμπεριφοράς οδήγησης.

Παρόλο που η αξιολόγηση της συμπεριφοράς οδήγησης έχει αποτελέσει αντικείμενο εκτεταμένης έρευνας και έχουν διεξαχθεί αρκετές μελέτες σχετικά με τις μεθοδολογίες συλλογής και ανάλυσης δεδομένων οδηγικής συμπεριφοράς, λίγες μόνο έχουν επικεντρωθεί στην απαιτούμενη ποσότητα δεδομένων που πρέπει να συλλέγονται. Σκοπός, λοιπόν, της παρούσας διπλωματικής εργασίας είναι η εκτίμηση του εύρους δεδομένων οδήγησης που απαιτείται να συλλέγονται για κάθε οδηγό για τον ακριβή προσδιορισμό της οδηγικής του συμπεριφοράς.

Για τη συλλογή δεδομένων χρησιμοποιήθηκαν έξυπνα κινητά τηλέφωνα (smartphones), τα οποία είχαν εγκατεστημένη μια εφαρμογή καταγραφής των δυναμικών και κινηματικών δεδομένων του οχήματος. Έτσι, αυτά τα δεδομένα φυσικής οδήγησης μπόρεσαν να δώσουν επαρκείς πληροφορίες για τα χαρακτηριστικά της συμπεριφοράς του κάθε οδηγού (π.χ. συχνότητα απότομων συμβάντων, απόσπαση της προσοχής κλπ.). Η συλλογή των δεδομένων έγινε στα πλαίσια προηγούμενων ερευνητικών εργασιών.

Επόμενο βήμα ήταν η αναζήτηση της κατάλληλης μεθόδου για την επεξεργασία και ανάλυση των δεδομένων. Για τον σκοπό αυτό εφαρμόστηκαν οι μέθοδοι Σωρευτικών Αθροισμάτων, Μεταβλητότητας Συμπεριφοράς και Κινητών Μέσων Όρων. Επιπλέον, για τον καθορισμό των παραμέτρων -

ορίων που θα καθόριζαν τις συνθήκες σύγκλισης / σταθεροποίησης της οδηγικής συμπεριφοράς, χρησιμοποιήθηκαν απλά στατιστικά όρια, με τα οποία ελέγχονταν η συμμόρφωση των κριτηρίων σύγκλισης που είχαν τεθεί.

Από τα αποτελέσματα της ανάλυσης προέκυψε ότι δεν υπάρχει ένα διακριτό χρονικό σημείο, κοινό για όλους τους οδηγούς ή/και όλα τα χαρακτηριστικά, από το οποίο και μετά η οδηγική συμπεριφορά να σταθεροποιείται. Διαφορετικοί παράγοντες φαίνεται να επηρεάζουν την σύγκλιση της οδηγικής συμπεριφοράς, με σημαντικότερους την επιθετικότητα των οδηγών συνολικά αλλά και ως προς το εκάστοτε χαρακτηριστικό που μελετάται, η μέση χρονική διάρκεια των ταξιδιών που εξετάζονται και η συνολική μεταβλητότητα (ή αντίθετα, σταθερότητα) της συμπεριφοράς του οδηγού από ταξίδι σε ταξίδι.

Μελλοντικά, το θέμα που εξετάστηκε στη διπλωματική εργασία θα μπορούσε να επεκταθεί εφαρμόζοντας διαφορετικές μεθόδους στατιστικής ανάλυσης όπως για παράδειγμα η ομαδοποίηση των οδηγών ανάλογα με τη συμπεριφορά τους χρησιμοποιώντας ομαδοποίηση K-μέσων ή και παρατηρώντας τις ίδιες μεταβλητές σε μεγαλύτερο δείγμα οδηγών. Ενδιαφέρον θα παρουσίαζε και μία ανάλυση η οποία θα στηριζόταν σε περισσότερα δημογραφικά χαρακτηριστικά των οδηγών, όπως το φύλλο, η ηλικία, η εμπειρία ή σε διαφορετικές συνθήκες κυκλοφορίας και διαφορετικά οδικά περιβάλλοντα.

Περιεχόμενα

1.	Εισαγωγή.....	17
1.1	Γενικά.....	17
1.1.1	Εκτίμηση της Οδηγικής Συμπεριφοράς.....	20
1.2	Σκοπός της Διπλωματικής Εργασίας.....	21
1.3	Μεθοδολογία	23
1.4	Δομή.....	25
2.	Βιβλιογραφική Ανασκόπηση	28
2.1	Εισαγωγή	28
2.2	Μέθοδοι Καταγραφής και Αξιολόγησης της Οδηγικής Συμπεριφοράς	28
2.2.1	Αυτοαξιολόγηση – Έρευνες με ερωτηματολόγια.....	29
2.2.2	Σύστημα Καταγραφής Δεδομένων εντός του οχήματος	32
2.2.3	Έξυπνα κινητά για την παρακολούθηση της οδηγικής συμπεριφοράς...	35
2.3	Μέθοδοι Ανάλυσης των δεδομένων οδηγικής συμπεριφοράς	39
2.3.1	Μέθοδοι για την Αξιολόγηση – Κατηγοριοποίηση της Οδηγικής Συμπεριφοράς	40
2.3.2	Μέθοδοι Ποσοτικοποίησης Δεδομένων	42
2.4	Γενικά Συμπεράσματα Βιβλιογραφικής Ανασκόπησης	44
3.	Μεθοδολογική Προσέγγιση	47

3.1	Ροή Εργασιών	47
3.1.1	Προκαταρκτική Επεξεργασία Δεδομένων	48
3.1.2	Σωρευτικά Αθροίσματα Οδηγικών Παραμέτρων	49
3.1.3	Μέθοδος Μεταβλητότητας Συμπεριφοράς	49
3.1.4	Κινητός Μέσος Όρος (MovingAverage)	51
3.2	Θεωρητικό Υπόβαθρο	53
3.2.1	Βασικές Έννοιες Χρονοσειρών	53
3.2.2	Μέθοδος Σωρευτικών Αθροισμάτων (CuSum)	57
3.2.3	Μεταβλητότητα Συμπεριφοράς	61
3.2.4	Κινητοί Μέσοι Όροι	64
3.2.5	Τοπικά Ακρότατα	69
4.	Συλλογή και Επεξεργασία Στοιχείων	73
4.1	Εισαγωγή	73
4.2	Συλλογή Στοιχείων	73
4.2.1	Μεθοδολογία καταγραφής, συλλογής, επεξεργασίας και οπτικοποίησης δεδομένων	74
4.3	Προκαταρκτική Ανάλυση Στοιχείων	90
5.	Ανάλυση και Αποτελέσματα	102
5.1	Εισαγωγή	102

5.2	Προετοιμασία των Δεδομένων.....	102
5.3	Διερεύνηση Μεταβλητότητας – Σύγκλισης των χαρακτηριστικών οδηγικής συμπεριφοράς.....	103
5.3.1	Περιγραφή της εισαγωγής των δεδομένων εισόδου στα Υπολογιστικά Φύλλα Excel	104
5.3.2	Περιγραφή της εξαγωγής των δεδομένων εξόδου από τα Υπολογιστικά Φύλλα Excel	105
5.4	Παρουσίαση Αποτελεσμάτων Ανάλυσης.....	110
6.	Συμπεράσματα.....	118
6.1	Σύνοψη Μεθοδολογίας και Αποτελεσμάτων	120
6.2	Βασικά Συμπεράσματα.....	121
6.3	Διατύπωση Προτάσεων.....	124
7.	Βιβλιογραφία.....	128
	ΠΑΡΑΡΤΗΜΑ	133

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 4.1: Επεξήγηση Μεταβλητών Βάσης Δεδομένων.....	83
Πίνακας 4.2: Επεξήγηση Νέων Μεταβλητών που προστέθηκαν στη Βάση Δεδομένων	88
Πίνακας 5.1: Διερεύνηση αργής ή γρήγορης σύγκλισης των χαρακτηριστικών οδηγικής συμπεριφοράς που μελετώνται με βάση τον ελάχιστο απαιτούμενο αριθμό ταξιδιών.....	114
Πίνακας 5.2: Όρια επιθετικότητας και μεταβλητότητας οδηγών όπως προέκυψαν από την ανάλυση.....	115
Πίνακας 5.3: Συγκεντρωτικός Πίνακας ποσοστών των οδηγών για το κρίσιμο χαρακτηριστικό τους, για κάθε διάρκεια	118
Πίνακας 0.1: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να συλλεχθούν για να επιτευχθεί σύγκλιση (Ταξίδια μέσης διάρκειας < 5 min.).....	134
Πίνακας 0.2: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να συλλεχθούν για να επιτευχθεί σύγκλιση (Ταξίδια μέσης διάρκειας 10 min.).....	135
Πίνακας 0.3: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να συλλεχθούν για να επιτευχθεί σύγκλιση (Ταξίδια μέσης διάρκειας 20 min.).....	136
Πίνακας 0.4: Μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδηγικής συμπεριφοράς (Ταξίδια μέσης χρονικής διάρκειας < 5 min.).....	139
Πίνακας 0.5: Μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδηγικής συμπεριφοράς (Ταξίδια μέσης χρονικής διάρκειας 10 min.).....	140
Πίνακας 0.6: Μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδηγικής συμπεριφοράς (Ταξίδια μέσης χρονικής διάρκειας 20 min.).....	141

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 3.1: Παράδειγμα χρονοσειράς, όπου παρατηρείται και η τάση, η εποχικότητα και τα κατάλοιπα (Πηγή: DigitalOcean).....	54
Εικόνα 3.2: Σύγκριση στάσιμης (πάνω) και μη-στάσιμης (κάτω) χρονοσειράς (Πηγή: Wikipedia).....	55
Εικόνα 3.3: Παράδειγμα Λευκού Θορύβου (Πηγή: ResearchGate).....	57
Εικόνα 3.4: Παράδειγμα Διαγράμματος Shewhart (Πηγή: Analyze-it).....	58
Εικόνα 3.5: Μετατόπιση μέσου κατανομής	59
Εικόνα 3.6: Παράδειγμα υψηλής και χαμηλής μεταβλητότητας.....	63
Εικόνα 3.7: Παράδειγμα μείωσης των συντελεστών βάρους για WMA (Πηγή: Wikipedia).....	67
Εικόνα 3.8: Παράδειγμα μείωσης των συντελεστών βάρους για EMA (Πηγή: Wikipedia).....	69
Εικόνα 3.9: Απόσπασμα χρονοσειράς γεγονότων απότομων επιταχύνσεων ενός τυχαίου χρήστη για ταξίδια μέσης διάρκειας 10 λεπτών	69
Εικόνα 3.10: Παράδειγμα τοπικών και ολικών ακροτάτων (Πηγή: Wikipedia).....	70
Εικόνα 4.1 Διάγραμμα διαχείρισης δεδομένων (Πηγή: OSeven)	74
Εικόνα 4.2: Σύστημα ροής δεδομένων (Πηγή: OSeven).....	76
Εικόνα 4.3: Yaw, Pitch, Roll (Πηγή: OSeven)	78
Εικόνα 4.4: Δείκτες κινδύνου οδήγησης (Πηγή: OSeven)	80

Εικόνα 4.5: Εφαρμογή στο κινητό και διαδικτυακή πύλη (Πηγή: OSeven).....	82
Εικόνα 5.1: Πεδίο εισαγωγής δεδομένων εισόδου στο αρχείο "RESULTS"	104
Εικόνα 5.2: Παράδειγμα μορφής δεδομένων εισόδου, όπως λαμβάνονται από τις υποβάσεις.....	105
Εικόνα 5.3: Καρτέλλα "HA" - Τμήμα ελέγχων για τη σύγκλιση του $cs_ha / cs_totaldist$	107
Εικόνα 5.4: Καρτέλλα "HA" - Τμήμα ελέγχων για τη σύγκλιση του μεγέθους μεταβλητότητας του $cs_ha / cs_totaldist$ ($Volatility[cs_ha / cs_totaldist]$) ...	107
Εικόνα 5.5: Πεδίο εισαγωγής τιμών ορίων και παραμέτρων ελέγχων.....	107

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1.1: Αίτια θανατηφόρων τροχαίων ατυχημάτων στην Ελλάδα για το 2017 (Πηγή: Ελληνική Αστυνομία, 2018)	18
Διάγραμμα 1.2: Ροή Αντιμετώπισης του Στόχου της Διπλωματικής Εργασίας.....	24
Διάγραμμα 3.1: Σχηματική Απεικόνιση Μεθοδολογικής Προσέγγισης	47
Διάγραμμα 4.1: Συνολική Καθαρή Διάρκεια Οδήγησης για κάθε οδηγό (σε ώρες) ...	91
Διάγραμμα 4.2: Συνολική Διανυθείσα Απόσταση για κάθε οδηγό (σε χιλιόμετρα)	91
Διάγραμμα 4.3: Αριθμός γεγονότων απότομων επιταχύνσεων ανά 100 km οδήγησης για κάθε οδηγό	92
Διάγραμμα 4.4: Αριθμός γεγονότων απότομων επιβραδύνσεων ανά 100 km οδήγησης για κάθε οδηγό	93
Διάγραμμα 4.5: Πόσοστό (%) του Συνολικού Καθαρού Χρόνου Οδήγησης κατά το οποίο χρησιμοποιήθηκε το κινητό τηλέφωνο, για κάθε οδηγό	94
Διάγραμμα 4.6: Ποσοστό (%) του Συνολικού Καθαρού Χρόνου οδήγησης κατά το οποίο υπήρξε υπέρβαση του ορίου ταχύτητας, για κάθε οδηγό.....	94
Διάγραμμα 4.7: Ιστόγραμμα Συχνοτήτων του αριθμού των Απότομων Επιταχύνσεων για κάθε οδηγό	95
Διάγραμμα 4.8: Ιστόγραμμα Συχνοτήτων του αριθμού των Απότομων Φρεναρισμάτων για κάθε οδηγό	95
Διάγραμμα 4.9: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων Απότομων Επιταχύνσεων προς τα σωρευτικά αθροιζόμενα km οδήγησης (κατακόρυφος άξονας) ως προς τη συνολική απόσταση οδήγησης (km) για κάθε οδηγό.....	96

Διάγραμμα 4.10: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων Απότομων Φρεναρισμάτων προς τα σωρευτικά αθροιζόμενα km οδήγησης (κατακόρυφος άξονας) ως προς τη συνολική απόσταση οδήγησης (km) για κάθε οδηγό.....97

Διάγραμμα 4.11: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων ωρών υπέρβασης του ορίου ταχύτητας προς τη σωρευτικά αθροιζόμενη διάρκεια οδήγησης (κατακόρυφος άξονας – hr/hr(%)) ως προς τη συνολική διάρκεια οδήγησης (hr) για κάθε οδηγό98

Διάγραμμα 4.12: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων ωρών χρήσης κινητού τηλεφώνου προς τη σωρευτικά αθροιζόμενη διάρκεια οδήγησης (κατακόρυφος άξονας – hr/hr(%)) ως προς τη συνολική διάρκεια οδήγησης (hr) για κάθε οδηγό99

Διάγραμμα 5.1: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του μεγέθους απότομων επιταχύνσεων ανά χιλιόμετρο και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες.....112

Διάγραμμα 5.2: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς τον αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο.....116

Διάγραμμα 0.1: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του μεγέθους απότομων επιταχύνσεων ανά χιλιόμετρο και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες.....137

Διάγραμμα 0.2: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του μεγέθους απότομων φρεναρισμάτων ανά χιλιόμετρο και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες.....137

Διάγραμμα 0.3: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του ποσοστού χρήσης κινητού τηλεφώνου και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες.....	138
Διάγραμμα 0.4: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του ποσοστού του χρόνου υπέρβασης ορίου ταχύτητας και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες.....	138
Διάγραμμα 0.5: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς τον αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο.....	142
Διάγραμμα 0.6: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς τον αριθμό απότομων φρεναρισμάτων ανά χιλιόμετρο	142
Διάγραμμα 0.7: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς το ποσοστό του χρόνου χρήσης κινητού τηλεφώνου κατά την οδήγηση.....	143
Διάγραμμα 0.8: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς το ποσοστό του χρόνου υπέρβασης του ορίου ταχύτητας κατά την οδήγηση	143

1. ΕΙΣΑΓΩΓΗ

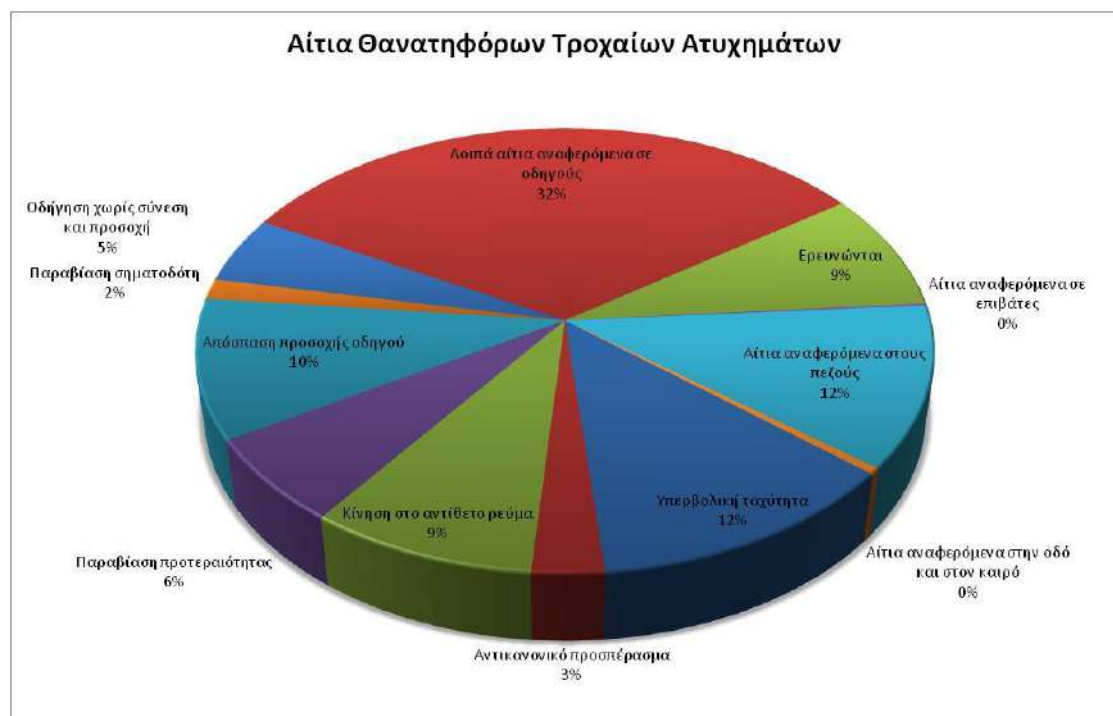
1.1 Γενικά

Οι παράγοντες κινδύνου που επηρεάζουν την πιθανότητα συμμετοχής σε τροχαίο ατύχημα είναι πολλοί και σημαντικός αριθμός αυτών έχει εντοπιστεί στη βιβλιογραφία. Μεταξύ άλλων, οι σημαντικότεροι παράγοντες κινδύνου που αναγνωρίζονται στη βιβλιογραφία (WHO 2015) είναι οι ανθρώπινοι παράγοντες όπως η ταχύτητα, η απόσπαση της προσοχής κατά την οδήγηση, η οδήγηση υπό την επήρεια αλκοόλ και άλλων ψυχοδραστικών ουσιών κλπ. Οι ανθρώπινοι παράγοντες θεωρούνται μία από τις κύριες αιτίες θανάτων από οδικές μεταφορές και τραυματισμούς κάθε χρόνο και ως εκ τούτου είναι εξαιρετικά σημαντικό να μελετηθεί πώς αυτοί οι παράγοντες μπορούν να επηρεάσουν τον κυκλοφοριακό κίνδυνο. Οι σημαντικότεροι παράγοντες που επιδρούν στην οδική ασφάλεια είναι τρεις, και παρουσιάζονται παρακάτω κατά σειρά αυξανόμενης σπουδαιότητας: **το όχημα** (τεχνική διαμόρφωση και εξοπλισμός, ενεργητική και παθητική ασφάλεια), **η οδός με το περιβάλλον της** (γεωμετρικά χαρακτηριστικά, κατασκευαστική διαμόρφωση, επίπεδο συντήρησης, εξοπλισμός, τοπικές κυκλοφοριακές ρυθμίσεις, χαρακτηριστικά της κυκλοφορίας, καιρικές συνθήκες), και **ο χρήστης της οδού** (εμπειρία, ψυχικές και κοινωνικές ιδιαιτερότητες καθώς επίσης πρότυπα συμπεριφοράς).

Σε πολλές χώρες, συμπεριλαμβανομένων των Ηνωμένων Πολιτειών, πολλοί ερευνητές μελέτησαν το ρόλο των τριών παραγόντων: "Άνθρωπος", "Οδός και Περιβάλλον" και "Όχημα" στη σοβαρότητα οδικών ατυχημάτων. Σύμφωνα με τα αποτελέσματα και τις διαπιστώσεις αυτών των μελετών, το 93% όλων των οδικών ατυχημάτων υπεύθυνος υπήρξε αποκλειστικά ο άνθρωπος ή τα οδικά ατυχήματα οφείλονταν σ' αυτόν σε συνδυασμό και με άλλους παράγοντες (Άνθρώπινος παράγοντας), ενώ το 34% όλων των οδικών ατυχημάτων οφείλεται στον παράγοντα οδοστρώματος. Ακολούθησε ο ρόλος του παράγοντα του οχήματος με τη συμβολή του 12% (Pakgothar et al., 2011). Κατά συνέπεια ο οδηγός δεν είναι μόνο ο παραλήπτης των πληροφοριών και εκείνος που παίρνει τις αποφάσεις, αλλά ο κύριος παράγοντας στα

ατυχήματα, καθώς τα περισσότερα από αυτά αποδίδονται στα ανθρώπινα λάθη, τα οποία οφείλονται κατά βάση στην αδύναμη συνειδητοποίηση των κινδύνων του οδηγού.

Στο Διάγραμμα 1.1 παρουσιάζονται πιο αναλυτικά οι κατηγορίες των παραγόντων που οδηγούν σε θανατηφόρα ατυχήματα στην Ελλάδα και οφείλονται σε ανθρώπινη υπαιτιότητα, με τα ποσοστά τους για το 2017. Όπως διαπιστώνουμε, ο ανθρώπινος παράγοντας ευθύνεται σχεδόν για το σύνολο των θανατηφόρων τροχαίων ατυχημάτων και επομένως κρίνεται απαραίτητη η συλλογή κατάλληλων και επαρκών στοιχείων, καθώς επίσης και η ανάλυσή τους για την μελέτη και την κατανόηση της οδηγικής συμπεριφοράς.



Διάγραμμα 1.1: Αίτια θανατηφόρων τροχαίων ατυχημάτων στην Ελλάδα για το 2017 (Πηγή: Ελληνική Αστυνομία, 2018)

Η βιβλιογραφία έχει δείξει ότι η **χρήση του κινητού τηλεφώνου** επηρεάζει σημαντικά την συμπεριφορά οδήγησης επειδή κάνει τους οδηγούς να οδηγούν πιο αργά με περισσότερες μεταβολές στην ταχύτητα, έχουν μεγαλύτερη διακύμανση στη θέση του πεντάλ γκαζιού και αναφέρουν υψηλότερο επίπεδο φόρτου εργασίας άσχετα από το επίπεδο δυσκολίας της

συνομιλίας. Οι οδηγοί επίσης τείνουν να επιλέγουν μεγαλύτερες αποστάσεις μεταξύ των οχημάτων και οι χρόνοι αντίδρασης τους αυξάνονται σημαντικά (Md Mazharul και Washington, 2015).

Η **οδήγηση πάνω από το όριο ταχύτητας** είναι ένας άλλος πολύ σημαντικός παράγοντας που επηρεάζει την πιθανότητα ατυχήματος (π.χ. μείωση της απόστασης αντίδρασης, απώλεια ελέγχου) και τη δύναμη πρόσκρουσης. Σύμφωνα με τον (ΟΟΣΑ, 2006), η οδήγηση πάνω από το όριο ταχύτητας αποτελεί παράγοντα που συμβάλλει στο 10% των συνολικών ατυχημάτων και πάνω από 30% σε θανατηφόρα ατυχήματα. Σύμφωνα με το Nilsson (1982), η πιθανότητα σύγκρουσης που περιλαμβάνει τραυματισμό είναι ανάλογη με το τετράγωνο της ταχύτητας, η πιθανότητα σοβαρής σύγκρουσης είναι ανάλογη προς τον κύβο της ταχύτητας και η πιθανότητα μιας θανατηφόρας σύγκρουσης σχετίζεται με την τέταρτη δύναμη της ταχύτητας.

Τέλος, τα **απότομα γεγονότα** όπως η **επιτάχυνση, το φρενάρισμα και οι στροφές** είναι τρεις σημαντικοί δείκτες για την αξιολόγηση της επικινδυνότητας οδήγησης (Bonsall et al., 2005), ειδικά για την αξιολόγηση της επιθετικότητας της οδήγησης. Αυτά τα χαρακτηριστικά συσχετίζονται έντονα με την επικίνδυνη απόσταση από τα γειτονικά οχήματα, τα παρολίγον ατυχήματα, την έλλειψη συγκέντρωσης, τον αυξημένο χρόνο αντίδρασης, την κακή κρίση οδήγησης ή την χαμηλή εμπειρία και συμμετοχή σε καταστάσεις υψηλού κινδύνου. Η συσχέτιση μεταξύ των εκδηλώσεων HA (HarshAcceleration) και HB (HarshBraking) με την επικίνδυνη οδήγηση έχει επισημανθεί στις επιστημονικές εργασίες που δημοσιεύθηκαν από τους Bonsall et al., (2005), Tselentis et al., (2017) και Mantouka et al. (2017) και έχει αναγνωριστεί ευρέως από την ασφαλιστική και την τηλεματική βιομηχανία.

Κατά συνέπεια, η καταγραφή και η αξιολόγηση τη συμπεριφοράς οδήγησης αποτελεί θέμα μείζονος σημασίας, τόσο για την βαθύτερη κατανόηση και πρόληψη των παραγόντων εκείνων που οδηγούν στα οδικά ατυχήματα, όσο και για την συνολική βελτίωση της οδηγικής συμπεριφοράς και παιδείας.

1.1.1 Εκτίμηση της Οδηγικής Συμπεριφοράς

Οι εκτιμήσεις οδηγικής συμπεριφοράς περιλαμβάνουν, γενικά, τόσο αξιολογήσεις εκτός δρόμου (off-road), όσο και πειράματα σε προσομοιωτές ή φυσικές καταστάσεις οδήγησης (Stav, 2004).

Η αξιολόγηση της οδήγησης **εκτός δρόμου** περιλαμβάνει έλεγχο ιστορικού οδήγησης, έλεγχο οράσεως και φυσικής κατάστασης, αξιολόγηση των γνωστικών και αντιληπτικών ικανοτήτων και εξέταση της γνώσης των ισχυόντων νόμων και κανόνων οδικής κυκλοφορίας. Η αξιολόγηση «off-road» βασίζεται σε ηλεκτρονικές αξιολογήσεις και αξιολογήσεις τύπου «μολύβι και χαρτί», ερωτηματολόγια αυτοαξιολόγησης και συνεντεύξεις. Παρόλο που αυτές οι εκτιμήσεις κοστίζουν σχετικά λίγα, υπόκεινται σε πιθανές προκαταλήψεις (Podsakoff and Organ, 1986, Maycock et al., 1991). Διεξάγονται σε κλινικό περιβάλλον και όχι στις πραγματικές συνθήκες κυκλοφορίας, γεγονός που δημιουργεί αμφιβολίες ως προς την αξιοπιστία τους ώστε να αντικατοπτρίζουν την πραγματική οδηγική συμπεριφορά στον πραγματικό κόσμο.

Οι αξιολογήσεις που πραγματοποιούνται σε **φυσικές καταστάσεις οδήγησης** μπορούν, αντιθέτως, να πραγματοποιηθούν στο αυτοκίνητο του ατόμου ή σε εκείνο του οδηγού αξιολόγησης για να μετρηθεί η απόδοση οδήγησης σε σχέση με τα πρότυπα της αρχής αδειοδότησης και να αξιολογηθεί η ικανότητα οδήγησης με ασφάλεια και νόμιμα. Παρόλο που η αξιολόγηση «πίσω από το τιμόνι» θεωρείται ότι έχει λογική εγκυρότητα, δεν επιτυγχάνεται πάντα η απόκτηση πληροφοριών σχετικά με προσωπικά χαρακτηριστικά οδήγησης, καθώς είναι δύσκολο να παρατηρηθούν άμεσα κατά τη διάρκεια της δοκιμής (Classen et al., 2010). Αυτοί οι περιορισμοί υποδηλώνουν την ανάγκη για άλλες, πιο αξιόπιστες πηγές λεπτομερών και αντικειμενικών πληροφοριών σχετικά με την οδήγηση σε πραγματικό χρόνο των ατόμων, προκειμένου να εκτιμηθεί η συμπεριφορά οδήγησης όταν αντιμετωπίζονται διάφορες καταστάσεις κυκλοφορίας και να παρακολουθούνται οι παραβιάσεις του νόμου.

Με τις πρόσφατες εξελίξεις στις τεχνολογίες παρακολούθησης, έχει δημιουργηθεί μια νέα γενιά μεθόδων συλλογής δεδομένων για την συνεχή εκτίμηση της πραγματικής συμπεριφοράς οδήγησης με μεγάλη λεπτομέρεια και τη δυνατότητα εφαρμογής μεγάλης κλίμακας. Οι **συσσκευές καταγραφής δεδομένων εντός οχήματος (IVDR)** είναι ηλεκτρονικές συσκευές παρακολούθησης που έχουν εγκατασταθεί στο όχημα και για πρώτη φορά καθιστούν δυνατή την ακριβή καταγραφή δεδομένων οδήγησης για μεγάλο χρονικό διάστημα για τη μελέτη των τρόπων οδήγησης (Shiftan and Toledo, 2008). Η συλλογή δεδομένων των συμβάντων για πολλές ώρες οδήγησης μπορεί να βοηθήσει στη συναρμολόγηση ενός προφίλ οδήγησης για τους οδηγούς αξιολογώντας την συμπεριφορά οδήγησης μέσω του ορισμού των συμβάντων (π.χ., απότομο φρενάρισμα, επιτάχυνση, απότομη στροφή και ταχύτητα) σε σχέση με αυτές τις μετρήσεις (Albert et al. 2011).

Σημειώνεται ότι οι προηγούμενες μελέτες δείχνουν ότι ανεξάρτητα από τον τύπο της τεχνολογίας μέτρησης, η οδήγηση πάνω από το όριο ταχύτητας και η απότομη πέδηση συνδέονται με υψηλότερα ποσοστά ατυχημάτων (Klauer et al., 2008, Simons-Morton et al., 2009, Kloeden κ.ά. van Schagen, 2006). Η παρακολούθηση της τάσης του κάθε οδηγού να επιδοθεί σε τέτοιες συμπεριφορές επιτρέπει στην τεχνολογία να υπολογίζει μια βαθμολογία κινδύνου για τον εν λόγω οδηγό (RoSPA, 2013).

1.2 Σκοπός της Διπλωματικής Εργασίας

Όλες οι παραπάνω τεχνικές συλλογής στοιχείων για την οδηγική συμπεριφορά ενισχύουν την χρησιμότητα των προχωρημένων συστημάτων συλλογής δεδομένων με κινητά τηλέφωνα. Παρ'όλα αυτά, **δεν προσδιορίζεται ακόμη το ακριβές μέγεθος των δεδομένων οδήγησης που πρέπει να συλλέγονται και να αξιολογούνται ώστε να προσδιοριστεί με ικανοποιητική ακρίβεια η συμπεριφορά οδήγησης.** Δεδομένου ότι τόσο τα μικρά όσο και τα μεγάλα δείγματα δεδομένων ενέχουν τον κίνδυνο να οδηγήσουν σε αμφισβητήσιμα αποτελέσματα, αποκτώντας ένα

δείγμα είτε μεροληπτικό είτε υπολογιστικά δαπανηρό για να αναλυθεί, είναι σημαντικό να διερευνηθεί πόσα δεδομένα οδήγησης πρέπει να καταγράφονται από κάθε συμμετέχοντα στο πείραμα. **Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μεθοδολογίας για την εκτίμηση του εύρους των δεδομένων οδήγησης που απαιτείται να συλλέγονται για κάθε οδηγό για τον προσδιορισμό της οδηγικής του συμπεριφοράς.**

Πιο συγκεκριμένα, επιδιώκεται να καθοριστεί το χρονικό διάστημα της οδήγησης ή ο αριθμός των ταξιδιών που είναι απαραίτητο να καταγραφούν, προκειμένου να αποκτηθεί μια ολοκληρωμένη εικόνα για κάθε οδηγό, όπου ο ρυθμός των οδηγικών χαρακτηριστικών (π.χ. ανά χιλιόμετρο ή ανά λεπτό) συγκλίνει σε μια σταθερή τιμή. Εξετάζεται δηλαδή αν υπάρχει συγκεκριμένο διακριτό χρονικό σημείο, πέρα από το οποίο τα χαρακτηριστικά της οδηγικής συμπεριφοράς δείχνουν να σταθεροποιούνται με το χρόνο, και επομένως η συλλογή περισσότερων δεδομένων δεν προσθέτει σημαντική επιπλέον γνώση.

Η βάση αυτού του πλαισίου είναι ένα καινοτόμο σύστημα συλλογής δεδομένων που καταγράφει συνεχώς τις αναλύσεις συμπεριφοράς οδήγησης κάθε συμμετέχοντα σε πραγματικό χρόνο, χρησιμοποιώντας αισθητήρες συσκευών έξυπνων κινητών τηλεφώνων (smartphones). Αναλύεται ένας μεγάλος αριθμός δεδομένων 68 οδηγών με διάφορα στατιστικά εργαλεία για να προσδιοριστεί ο ελάχιστος χρόνος παρατήρησης κάθε οδηγού και η δυνατότητα ομαδοποίησης οδηγών με κριτήριο την επιθετικότητα της συμπεριφοράς οδήγησής τους.

Τα αποτελέσματα της ανάλυσης και τα συμπεράσματα της εργασίας αναμένεται να αποφέρουν **πολλαπλά και σημαντικά οφέλη**, εφόσον θα είναι δυνατό να αξιοποιηθούν τόσο ατομικά από τους οδηγούς για την βελτίωση της οδηγικής τους συμπεριφοράς, όσο και από τις ασφαλιστικές εταιρείες και τις βιομηχανίες αυτοκινήτων για τη βελτίωση της παροχής υπηρεσιών προς τους πελάτες τους. Αξίζει όμως να αναφερθούν και τα κοινωνικά οφέλη, καθώς αυτή η μέθοδος θα βοηθήσει στην συνολική βελτίωση της συμπεριφοράς

οδήγησες και θα μειώσει την εκπομπή ρύπων, τον κορεσμό, και την κατανάλωση ενέργειας, αναβαθμίζοντας γενικά τα μεταφορικά συστήματα.

1.3 Μεθοδολογία

Αρχικά μετά την οριστικοποίηση του επιδιωκόμενου στόχου πραγματοποιείται **βιβλιογραφική ανασκόπηση** τόσο σε ελληνική όσο και σε διεθνή βιβλιογραφία. Στη φάση αυτή πραγματοποιείται αναζήτηση παρεμφερών ερευνών, επιστημονικών άρθρων καθώς και γενικών πληροφοριών σχετικά με το εξεταζόμενο αντικείμενο που μπορούν να φανούν χρήσιμες για τη συγκεκριμένη έρευνα. Μέσω των ερευνών αυτών καταβάλλεται προσπάθεια να αποκτηθεί μια σχετική εμπειρία στην επεξεργασία τέτοιων θεμάτων, καθώς επίσης και να αποφασιστεί η μέθοδος με βάση την οποία θα πραγματοποιηθεί η επεξεργασία των στοιχείων και να επιτευχθεί ο επιδιωκόμενος στόχος.

Μετά την διερεύνηση των στοιχείων που συγκεντρώθηκαν από τη βιβλιογραφική ανασκόπηση, σειρά έχει η εύρεση του τρόπου συλλογής στοιχείων. Εφαρμόζεται ένα καινοτόμο σύστημα συλλογής δεδομένων, μέσα από την εξατομικευμένη καταγραφή της συμπεριφοράς του οδηγού σε πραγματικό χρόνο χρησιμοποιώντας αισθητήρες έξυπνων κινητών τηλεφώνων. **Εξήντα οκτώ (68) οδηγοί συμμετείχαν στο πείραμα οδήγησης σε πραγματικές συνθήκες** (smartphone naturalistic driving experiment) σε χρονικό διάστημα 7 μηνών και δημιουργήθηκε μια μεγάλη βάση δεδομένων με χιλιάδες διαδρομές. Για τον σκοπό αυτό αξιοποιείται η ειδική πλατφόρμα της εταιρείας της συγκεκριμένης εφαρμογής σε έξυπνα κινητά τηλέφωνα (OSeven) για την αποτελεσματική μετάβαση από τη συλλογή στοιχείων στη διαδικασία ανάλυσης των δεδομένων.

Στη συνέχεια, ακολουθεί η επιλογή της μεθόδου στατιστικής επεξεργασίας των στοιχείων και η εισαγωγή της βάσης δεδομένων, που έχει προετοιμαστεί στο Microsoft Excel, στο ελεύθερο λογισμικό RStudio και στο Microsoft Excel.

Επόμενο στάδιο αποτελεί η **χρήση των διαφόρων στατιστικών εργαλείων και μεθόδων** όπως είναι οι μέθοδοι Χρονοσειρών, CUSUM και Μεταβλητότητας Συμπεριφοράς (Volatility) που συντελούν στην κατανόηση και την επεξεργασία της βάσης δεδομένων ώστε να επιτευχθούν οι στόχοι της παρούσας διπλωματικής εργασίας. Στα τελικά δεδομένα που χρησιμοποιούνται λαμβάνεται υπόψη και η διάρκεια οδήγησης και δίνεται ιδιαίτερη σημασία στην χρονολογική ταξινόμηση των δεδομένων προκειμένου να αποφευχθούν σφάλματα στις χρονοσειρές και στα χαρακτηριστικά ομοιογένειάς τους.

Μετά την **αξιολόγηση και ερμηνεία** των αποτελεσμάτων, εξάγονται τα αντίστοιχα συμπεράσματα για τον χρόνο της παρακολούθησης των οδηγών. Έτσι, προκύπτουν σημαντικές πληροφορίες για το υπό εξέταση πρόβλημα καθώς και διατυπώνονται αξιολογες προτάσεις για περαιτέρω έρευνα.



Διάγραμμα 1.2: Ροή Αντιμετώπισης του Στόχου της Διπλωματικής Εργασίας

1.4 Δομή

Το **πρώτο κεφάλαιο** αποτελεί την εισαγωγή και περιλαμβάνει μια γενική ανασκόπηση όπου παρουσιάζεται το πλαίσιο της Διπλωματικής Εργασίας και αναφέρονται τα δεδομένα όπως έχουν προκύψει ως σήμερα, όσον αφορά την οδηγική συμπεριφορά. Παρατίθενται κάποια βασικά στατιστικά στοιχεία και τέλος παρουσιάζεται η μεθοδολογία που ακολουθήθηκε, περιγράφοντας συνοπτικά τα βασικά στάδια της εκτέλεσης της παρούσας εργασίας.

Στο **δεύτερο κεφάλαιο** παρουσιάζονται τα αποτελέσματα της βιβλιογραφικής ανασκόπησης, τα οποία προέκυψαν από την αναζήτηση και την καταγραφή ερευνών με αντικείμενο την μελέτη και καταγραφή της οδηγικής συμπεριφοράς καθώς επίσης και τις μεθόδους συλλογής στοιχείων από πειράματα σε πραγματικές συνθήκες οδήγησης. Εξετάζονται εργασίες από Ελλάδα και εξωτερικό, έρευνες που έχουν δημοσιευθεί σε επιστημονικά περιοδικά, βιβλία, άρθρα, πρακτικά συνεδρίων. Το κεφάλαιο κλείνει με τη σύνοψη και κριτική αξιολόγηση των αποτελεσμάτων των ερευνών.

Στο **τρίτο κεφάλαιο** παρουσιάζεται το θεωρητικό υπόβαθρο στο οποίο στηρίχθηκε η ανάλυση των στοιχείων της οδηγικής συμπεριφοράς. Αρχικά, γίνεται μία σύντομη αναφορά στα βήματα που ακολουθούνται, για την επεξεργασία των δεδομένων στα ειδικά λογισμικά στατιστικής ανάλυσης. Στη συνέχεια επεξηγούνται οι λόγοι επιλογής των διαφόρων μαθηματικών και στατιστικών αναλύσεων που χρησιμοποιήθηκαν, και τέλος, περιγράφονται βασικές μαθηματικές και στατιστικές έννοιες για την καλύτερη κατανόηση και επεξήγηση των μεθόδων αυτών.

Το **τέταρτο κεφάλαιο** περιγράφει τη διαδικασία συλλογής και επεξεργασίας των στοιχείων. Αρχικά, περιγράφεται η μέθοδος συλλογής των δεδομένων μέσα από το πείραμα σε πραγματικές συνθήκες οδήγησης που πραγματοποιήθηκε. Στη συνέχεια παρατίθενται συγκεντρωτικά διαγράμματα τα οποία περιγράφουν τα χαρακτηριστικά του δείγματος, ακολουθούμενα από τον απαραίτητο σχολιασμό.

Στο **πέμπτο κεφάλαιο** περιλαμβάνεται η αναλυτική περιγραφή της μεθοδολογίας που εφαρμόστηκε ως την εξαγωγή των τελικών αποτελεσμάτων. Αρχικά περιγράφονται τα βήματα που ακολουθήθηκαν για την εφαρμογή της μεθοδολογίας και παρουσιάζεται η διαδικασία ανάπτυξης της ανάλυσης. Παρουσιάζονται τα δεδομένα εισόδου κι εξόδου, με ιδιαίτερη έμφαση στα αποτελέσματα και στην ερμηνεία τους, ενώ τα τελικά αποτελέσματα συνοδεύονται από την περιγραφή τους για την διευκόλυνση της κατανόησης τους.

Στο **6ο κεφάλαιο** διατυπώνονται τα κυριότερα συμπεράσματα, τα οποία προέκυψαν από την ανάλυση των στοιχείων της έρευνας. Στην αρχή του κεφαλαίου πραγματοποιείται μία σύνοψη των κυριότερων σημείων της παρούσας εργασίας, ενώ στη συνέχεια διατυπώνονται τα συμπεράσματα, τα οποία προέκυψαν. Στο τέλος του κεφαλαίου, διατυπώνονται προτάσεις, τόσο σχετικά με την αξιοποίηση των αποτελεσμάτων της παρούσας εργασίας, όσο και σχετικά με τη βελτίωση της οδικής ασφάλειας και περαιτέρω έρευνες, σχετικές με το αντικείμενο της εργασίας αυτής.

Στο **7ο κεφάλαιο** παρατίθεται ο κατάλογος των βιβλιογραφικών αναφορών. Το κεφάλαιο αυτό περιλαμβάνει αναφορές, που αφορούν τόσο σε έρευνες που παρουσιάστηκαν στα κεφάλαια της εισαγωγής και της βιβλιογραφικής ανασκόπησης, όσο και σε στατιστικές έννοιες και μεθόδους, που αναλύθηκαν στο θεωρητικό υπόβαθρο.

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 Εισαγωγή

Η συμπεριφορά του οδηγού και τα σφάλματα κατά τη διάρκεια της οδήγησης αποτελούν μια από τις σημαντικότερες αιτίες οδικών ατυχημάτων. Η οδηγική συμπεριφορά σχετίζεται με το χαρακτήρα του οδηγού και το κοινωνικοοικονομικό υπόβαθρο, ωστόσο επηρεάζεται και μέσω της εκπαίδευσης, της κατάρτισης, των εκστρατειών δημοσιότητας και της επαρκούς αστυνόμευσης. Επομένως, η κατανόηση και η επηρεασμός της συμπεριφοράς των οδηγών αποτελεί βασικό συστατικό για την βελτίωση της οδικής ασφάλειας.

Ένα από τα κύρια εμπόδια στην κατανόηση των σχέσεων μεταξύ των χαρακτηριστικών των οδηγών και της οδηγικής τους συμπεριφοράς είναι η έλλειψη αξιόπιστων εργαλείων για τη συλλογή λεπτομερών πληροφοριών σχετικά με τα άτομα, όπως το επίπεδο δεξιοτήτων τους και οι οδηγικές τους ικανότητες, καθώς επίσης και η παρακολούθηση και ερμηνεία της οδηγικής τους συμπεριφοράς, όπως καταγράφονται για παράδειγμα από προφίλ επιτάχυνσης και ταχύτητας. Συνεπώς, το πρόβλημα αυτό έχει απασχολήσει ιδιαίτερωσ την έρευνα γύρω από τα θέματα οδικής ασφάλειας, και πλήθος ερευνητικών εργασιών και πειραμάτων έχουν αφιερωθεί στην προσπάθεια καταγραφής και ανάλυσης της οδηγικής συμπεριφοράς.

2.2 Μέθοδοι Καταγραφής και Αξιολόγησης της Οδηγικής Συμπεριφοράς

Μεταξύ των πιο συνηθισμένων μεθοδολογιών που χρησιμοποιούνται είναι οι προσομοιωτές οδήγησης (Desmond et al., 1998), ερωτηματολόγια (Gerald et al., 1998) σε συνδυασμό με προσομοιωτές και πειράματα οδήγησης σε πραγματικές συνθήκες (Toledo et al., 2008), ενώ τα τελευταία χρόνια οι πιο συνηθισμένες μέθοδοι για την καταγραφή των χαρακτηριστικών της οδήγησης

είναι οι καταγραφείς εντός οχημάτων (Shichrur et al., 2014) και πιο πρόσφατα, η εκμετάλλευση των smartphones (Vlahogianni and Barmounakis ,2017).

2.2.1 Αυτοαξιολόγηση – Έρευνες με ερωτηματολόγια

Οι περισσότερες μελέτες που αξιολογούν την οδηγική συμπεριφορά επικεντρώνονται στην αυτο-εκτίμηση, τις αντιλήψεις και τις στάσεις των οδηγών απέναντι στο έργο της οδήγησης. Αυτές οι μελέτες βασίζονται συχνά σε απαντήσεις από ερωτηματολόγια, στα οποία οι ίδιοι οι οδηγοί καλούνται να αξιολογήσουν τη δική τους οδηγική συμπεριφορά, τις στάσεις και τις αντιλήψεις τους για την οδήγηση, όπως η ανάληψη κινδύνου και η υπακοή στον Κώδικα Οδικής Κυκλοφορίας, καθώς επίσης και να αναφέρουν το ιστορικό τους σχετικά με την οδική ασφάλεια (π.χ. εμπλοκή σε ατυχήματα, αναφορές της αστυνομίας κλπ.).

Αυτή η μέθοδος έχει πολλά σημαντικά πλεονεκτήματα. Το πιο σημαντικό είναι ότι μπορεί να χρησιμοποιηθεί για τη συλλογή μεγάλου όγκου δεδομένων σε σχετικά σύντομο χρονικό διάστημα και με χαμηλό κόστος. Ως αποτέλεσμα, αυτή η προσέγγιση έχει χρησιμοποιηθεί εκτεταμένα για ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένων των μελετών επιθετικής οδήγησης (Parker et al. 1998, Lajunen et al. 1998, Chliaoutakis et al. 2002, Ulleberg & Rundmo 2002, Iversen 2004, Maxwell et al. 2005), οι σχέσεις μεταξύ χρήσης αλκοόλ και ναρκωτικών και (Shinar 1995, Schechtman et al. 1999, Caetano & Clark 2000, Begg & Langley 2004), κοινωνικοοικονομικά χαρακτηριστικά που επηρεάζουν τη συμπεριφορά των οδηγών (Yagil 1998, Lourens et al. 1999, Golias & Karlaftis 2001, Boyce & Geller 2002) και άλλα.

Έχουν γίνει πολλές προσπάθειες να χρησιμοποιηθούν οι στάσεις και οι αντιλήψεις των οδηγών ως προς το έργο της οδήγησης για την πρόβλεψη της συμπεριφοράς των οδηγών. Η προσέγγιση υποθέτει ότι είναι δυνατόν να προσδιοριστούν ορισμένες στάσεις που συνδέονται με τη συμπεριφορά, καθώς επίσης και ότι οι στάσεις αυτές σχετίζονται αιτιωδώς με τη

συμπεριφορά. Οι περισσότερες μελέτες βρήκαν σχετικά χαμηλή συσχέτιση μεταξύ των δύο (Forward, 1997). Ωστόσο, αυτό το αποτέλεσμα οδήγησε μόνο σε αυξημένες προσπάθειες για την δημιουργία μιας ισχυρότερης σχέσης μεταξύ τα δύο.

Η **«Θεωρία της Λογικής Δράσης»** (Theory of Reasoned Action - **TRA**) αναπτύχθηκε από τους Fishbein και Ajzen (1975) οι οποίοι διαπίστωσαν ότι οι οδηγοί διαφέρουν μεταξύ τους λόγω της ετοιμότητάς τους να διαπράξουν κυκλοφοριακή παραβίαση του νόμου. Η θεωρία δηλώνει ότι η πρόθεση να διαπράξει το άτομο κάποια συγκεκριμένη συμπεριφορά μπορεί να προβλεφθεί από τις προσωπικές του πεποιθήσεις και στάσεις απέναντι σε αυτή τη συμπεριφορά.

Αργότερα αναπτύχθηκε η **«Θεωρία της Προγραμματισμένης Συμπεριφοράς»** (Theory of Planned Behavior - **TPB**) (Ajzen, 1985). Η TPB προβλέπει συμπεριφορές με βάση τις προθέσεις και τις στάσεις απέναντι σε μια πράξη και τον αντιληπτό έλεγχο συμπεριφοράς. Οι προθέσεις αναφέρονται στην προθυμία να δοκιμάσει το άτομο μια συγκεκριμένη συμπεριφορά, και θεωρεί το ισχυρότερο προγνωστικό αυτής. Οι στάσεις αναφέρονται στην ατομική εκτίμηση της εκτέλεσης μιας συγκεκριμένης συμπεριφοράς ενώ ο αντιληπτός έλεγχος συμπεριφοράς αναφέρεται στις πεποιθήσεις των ανθρώπων για την αποδοχή ή μη αποδοχή μιας συγκεκριμένης συμπεριφοράς (Aberg, 1997).

Η «Θεωρία της Λογικής Δράσης» εφαρμόστηκε σε μεγάλο αριθμό οδηγικών συμπεριφορών και διαπιστώθηκε ότι οι προβλέψεις του για την οδηγική συμπεριφορά ήταν σημαντικά ακριβείς (Parker et al., 1992, Yagil, 1998, Gordon and Hunt, 1998). Οι ερευνητές ασχολούνταν κυρίως με τις προθέσεις των ατόμων να διαπράττουν συμπεριφορές όπως η οδήγηση πάνω από το όριο ταχύτητας ή οδήγηση υπό την επήρεια αλκοόλ. Όσο για τη «Θεωρία της Προγραμματισμένης Συμπεριφοράς», εκτενής ανασκόπηση και παραδείγματα μπορούν να βρεθούν στο Ajzen (2006).

Παρά την κοινή αποδοχή και πρακτική χρήση της μεθόδου αυτοαξιολόγησης με ερωτηματολόγια, πολλοί ερευνητές εξέφρασαν αμφιβολίες σχετικά με την αξιοπιστία των αναφορών για τη μέτρηση της συμπεριφοράς των οδηγών, καθώς ενδέχεται να υπάρχουν διάφορες μεροληπτικές απαντήσεις.

Για παράδειγμα, οι Maycock et al. (1991) δείχνουν ότι ο αριθμός των αυτοαναφερόμενων τροχαίων ατυχημάτων ενδέχεται να είναι χαμηλότερος από τον πραγματικό αριθμό ακόμα και κατά 30%. Για τα παρ' ολίγον ατυχήματα και τις παραβιάσεις του Κ.Ο.Κ., το ποσοστό μη αναφοράς ήταν ακόμη υψηλότερο. Αυτή η ανεπαρκής αναφορά δεν είναι εύκολο να διορθωθεί από άλλες βάσεις δεδομένων που περιέχουν πληροφορίες για αυτοκινητιστικά ατυχήματα, όπως αυτές που διατηρούνται από την αστυνομία, τα νοσοκομεία και τις ασφαλιστικές εταιρείες καθώς και αυτές επίσης υποφέρουν από σημαντικές ελλείψεις στην αρχειοθέτηση (Hauer & Hakkert, 1988).

Οι Arthur et al. (2001) αξιολόγησαν επίσης τη συσχέτιση μεταξύ αρχειοθετημένων και αυτοαναφερόμενων τροχαίων ατυχημάτων και παραβάσεων του Κ.Ο.Κ. και βρήκαν χαμηλή συσχέτιση μεταξύ των δύο. Η έρευνα αναφέρει επίσης αρκετές άλλες έρευνες όπου διαπιστώθηκε σημαντική υπο-αναφορά ατυχημάτων και παραβάσεων του Κ.Ο.Κ.. Παραπλανητικές αναφορές εμφανίζονται επίσης σε άλλους τύπους ερωτήσεων. Οι περισσότεροι οδηγοί, και ειδικά οι άντρες, τείνουν να υπερεκτιμούν τις οδηγικές ικανότητές τους (McKenna et al., 1991). Σημαντικός αριθμός μεροληπτικών απαντήσεων παρατηρήθηκε όταν οι οδηγοί κλήθηκαν να υποδείξουν πόσο συχνά πραγματοποιούσαν μια συγκεκριμένη συμπεριφορά, με μεγαλύτερη μεροληψία να εμφανίζεται στις ανεπιθύμητες συμπεριφορές σε σύγκριση με τις επιθυμητές (Linderman & Verkasalo, 1995). Οι Lajunen και Summala (2003) διαπίστωσαν ότι αυτή η «κοινωνικά επιθυμητή μεροληψία» στην αυτοαναφερόμενη συμπεριφορά του οδηγού εξαρτάται από τον σχεδιασμό του ερωτηματολογίου και τις διαδικασίες της έρευνας.

2.2.2 Σύστημα Καταγραφής Δεδομένων εντός του οχήματος

Αυτοί οι περιορισμοί της αυτό-αναφοράς ως βάση για τη διερεύνηση της οδηγικής συμπεριφοράς σαφώς υποδεικνύουν την ανάγκη για άλλες, πιο αξιόπιστες πηγές πληροφοριών, που μπορούν είτε να χρησιμοποιηθούν συμπληρωματικά με τα ερωτηματολόγια, είτε να παρέχουν μόνες τους τα απαιτούμενα δεδομένα. Μια τέτοια κατηγορία εργαλείων που αναπτύχθηκε πρόσφατα είναι τα **Συστήματα Καταγραφής Δεδομένων εντός του οχήματος (IVDR)**. Τα IVDR εγκαθίστανται στο όχημα και παρέχουν λεπτομερείς πληροφορίες σχετικά με τη θέση, την ταχύτητα, την κάθετη και οριζόντια επιτάχυνση και τους ελιγμούς που εκτελούνται.

Η πρώτη εφαρμογή των καταγραφικών δεδομένων οχημάτων ήταν η συσκευή καταγραφής δεδομένων συμβάντων (Event Data Recorder - EDR). Το EDR είναι παρόμοιο με το "μαύρο κουτί" που χρησιμοποιείται στα αεροσκάφη και καταγράφει δεδομένα, όταν συμβάντα, όπως συγκρούσεις, συμβαίνουν αποθηκεύοντας τις πληροφορίες στη μονάδα. Αυτές οι πληροφορίες χρησιμοποιούνται αργότερα για να διερευνηθούν και να αναλυθούν οι συνθήκες που οδήγησαν στη σύγκρουση. Τα πρώτα πειράματα με τα EDR διεξήχθησαν από την NHTSA (National Highway Traffic Safety Administration - στις ΗΠΑ) στη δεκαετία του '70. Σήμερα τα EDR εγκαθίστανται ευρέως και χρησιμοποιούνται από κατασκευαστές οχημάτων, ασφαλιστικές εταιρείες, υπηρεσίες επιβολής του νόμου και ερευνητές. Ενώ αυτά τα εργαλεία είναι χρήσιμα για τη μελέτη του ίδιου του συμβάντος σύγκρουσης, τα δεδομένα που αποθηκεύουν περιορίζονται σε διάστημα μερικών δευτερολέπτων πριν από την εκδήλωση του ατυχήματος και δεν επαρκούν για μια πιο γενική μελέτη της συμπεριφοράς του οδηγού.

Πλέον η χρήση συσκευών καταγραφής έχει επεκταθεί στη μελέτη της συμπεριφοράς των οδηγών και σε καταστάσεις μη-σύγκρουσης. Το IVDR τέθηκε για πρώτη φορά στη βιομηχανία φορτηγών περίπου είκοσι χρόνια πριν. Ο εξοπλισμός προοριζόταν αρχικά για καθήκοντα διαχείρισης στόλου, όπως η δρομολόγηση και η παρακολούθηση. Από τότε έχουν προστεθεί πολλές άλλες λειτουργίες σε αυτές τις συσκευές, συμπεριλαμβανομένων και

ορισμένων που προσανατολίζονται στην παρακολούθηση και τη βελτίωση της συμπεριφοράς του οδηγού και την οδική ασφάλεια.

Στο πλαίσιο αυτό, χρησιμοποιήθηκαν IVDR ως μέρος σχεδίων **Ασφάλειας με βάση την Συμπεριφορά** (Behavioral - Based Safety - **BBS**) που εφαρμόζουν ένα σύνολο μεθόδων που έχει αποδειχθεί ότι βελτιώνουν τις επιδόσεις ασφαλείας. Αυτό επιτυγχάνεται εμπλέκοντας τους εργαζομένους στη διαδικασία βελτίωσης, διδάσκοντάς τους πώς να προσδιορίζουν τις κρίσιμες για την ασφάλεια συμπεριφορές και να πραγματοποιούν παρατηρήσεις για τη συλλογή δεδομένων, ενώ παράλληλα τους παρέχονται σχόλια για να ενθαρρύνουν τη βελτίωση και τη χρήση των δεδομένων που συγκεντρώθηκαν για να στοχεύσουν προς θετικές αλλαγές (Krause et al., 1999). Οι Krause et al. (1999) κατέληξαν στο συμπέρασμα ότι η BBS έχει σημαντικές δυνατότητες βελτίωσης της ασφάλειας των φορτηγών, ειδικά εάν χρησιμοποιούνται νέες τεχνολογίες όπως το IVDR. Αρκετές μελέτες έδειξαν ότι, όταν παρέχονται εβδομαδιαίες αναφορές απόδοσης ασφαλείας στους επιβλέποντες και τους διαχειριστές στόλων ως μέρος μιας υλοποίησης του BBS, αυτές έχουν θετικό αντίκτυπο στην ασφάλεια (Grindle et al., 2000; Sulzer-Azaroff & Austin, 2000).

Εκτός όμως από τους διαχειριστές στόλων οχημάτων, και οι ίδιοι οι οδηγοί όταν λαμβάνουν ανατροφοδότηση (feedback) σχετικά με τις αποδόσεις τους στην οδήγηση, μπορούν να βελτιώσουν την συμπεριφορά τους και συνεπώς την οδική ασφάλεια γενικότερα. Μελέτες που αξιολόγησαν αυτήν την επίδραση της ανατροφοδότησης στους οδηγούς σχετικά με την οδηγική συμπεριφορά τους βρήκαν γενικά σημαντικές βελτιώσεις ασφαλείας. Για παράδειγμα, οι Wouters και Bos (1997) πραγματοποίησαν ένα πείραμα με 840 οχήματα, 270 από τα οποία είχαν εξοπλιστεί με IVDR. Βρήκαν 20% μείωση των συντριβών αυτοκινήτων για τους οδηγούς που αντιμετώπιζαν τα δεδομένα που καταγράφηκαν από το IVDR.

Η επιτυχία του IVDR μπορεί να αποδοθεί εν μέρει στο ότι πρόκειται περί ενός αντικειμενικού συστήματος καταγραφής βασισμένου στην τεχνολογία. Χρησιμοποιώντας ομάδες εστίασης, ο Roetting et al. (2003) διαπίστωσε ότι οι οδηγοί είχαν θετική στάση απέναντι στην ανατροφοδότηση σχετικά με την

οδήγησή τους και ιδιαίτερα όταν προσφέρονται από μια νέα τεχνολογία. Από την άλλη πλευρά, ο Heinzmann & Schade (2003), ο οποίος διερεύνησε εάν η παρουσία μονάδων αποθήκευσης δεδομένων οδήγησης στα αυτοκίνητα νέων ανδρών οδηγών είχε μπορούσε να λειτουργήσει προληπτικά οδηγώντας σε μεγαλύτερη πειθαρχία και προσεκτική οδήγηση, έδειξε ότι η εγκατάσταση της μονάδας από μόνη της δεν είχε σημαντικές επιπτώσεις στη συμπεριφορά ή την εμφάνιση ατυχημάτων.

Η ερμηνεία των δεδομένων από το IVDR

Παρόλο που οι μελέτες αυτές δείχνουν σαφώς ότι η τεχνολογία IVDR μπορεί να έχει σημαντικό αντίκτυπο στην οδική ασφάλεια, υπάρχει περιορισμένη κατανόηση του ποια δεδομένα πρέπει να συλλέγονται και πώς πρέπει να ερμηνεύονται και να αναλύονται. Για παράδειγμα, δεν είναι ακόμη σαφές πώς τα δευτερεύοντα προφίλ ταχύτητας και επιτάχυνσης μεταφράζονται σε κίνδυνο ατυχήματος. Καθώς τα δεδομένα IVDR γίνονται πιο προσιτά, αυξάνεται η σημασία της σωστής ερμηνείας τους.

Κάποιες απόπειρες προς αυτήν την κατεύθυνση έχουν ξεκινήσει τα τελευταία χρόνια. Η μελέτη **Drive Atlanta** (Georgia Tech, 2002) επικεντρώνεται στη συλλογή δεδομένων IVDR από 1100 εξοπλισμένα οχήματα. Τα δεδομένα αποστέλλονται εβδομαδιαίως και περιλαμβάνουν πληροφορίες επιπέδου ταξιδιού, όπως η διανυθείσα απόσταση, η διάρκεια ταξιδιού και η επιλογή διαδρομής, καθώς και η ταχύτητα και η επιτάχυνση ανά δευτερόλεπτο. Το **TripSense** (TripSense, 2005) είναι ένα παρόμοιο πρόγραμμα που πραγματοποιήθηκε στη Μινεσότα. Στο επίπεδο ταξιδιού, τα δεδομένα που συλλέγονται περιλαμβάνουν την ώρα έναρξης και λήξης των ταξιδιών, την απόσταση και τη διάρκεια του ταξιδιού. Τα δεδομένα συμπεριφοράς οδηγού περιλαμβάνουν τις ταχύτητες σε διαστήματα 10 δευτερολέπτων και τους αριθμούς των επιθετικών γεγονότων πέδησης και των επιθετικών επιταχύνσεων επιτάχυνσης. Η επιθετική πέδηση και η επιτάχυνση ορίζονται (μάλλον αυθαίρετα), ως άνω των 7 mph ανά δευτερόλεπτο. Το πείραμα

διεξάγεται σε συνδυασμό με μια ασφαλιστική εταιρεία, η οποία επιτρέπει στους οδηγούς να δουν τις δικές τους αναφορές οδήγησης σε μια ειδική ιστοσελίδα και να συγκριθούν με άλλους οδηγούς.

Πρόσφατα στο Ισραήλ διεξήχθη μια έρευνα που περιελάμβανε 33 συμμετέχοντες των οποίων τα οχήματα ήταν εξοπλισμένα με συστήματα IVDR. Κατά τη διάρκεια της πρώτης φάσης του πειράματος οι οδηγοί δεν έλαβαν κανένα σχόλιο από το IVDR, ενώ στη δεύτερη φάση οι οδηγοί είχαν πρόσβαση σε μια προσωπική ιστοσελίδα που συνοψίζει τις πληροφορίες που καταγράφηκαν από το IVDR. Τα αποτελέσματα έδειξαν σημαντική συσχέτιση μεταξύ των ιστορικών καρτών καταγραφής των οδηγών και των μετρήσεων IVDR. Επιπλέον, τα αποτελέσματα έδειξαν ότι η αρχική έκθεση στο σύστημα είχε σημαντικές θετικές επιπτώσεις στη συμπεριφορά των οδηγών (Toledo & Lotan, 2006).

Το IVDR παρέχει πολύ πιο πλούσιες πληροφορίες σχετικά με τη συμπεριφορά των οδηγών, οι οποίες δεν ήταν διαθέσιμες παλαιότερα. Συνοπτικά, η τεχνολογία IVDR, η οποία παρακολουθεί την συμπεριφορά οδήγησης καταγράφοντας τις θέσεις, την ταχύτητα και την επιτάχυνση των οχημάτων, μπορεί να έχει σημαντικές θετικές επιπτώσεις στην οδική ασφάλεια. Ωστόσο, προκειμένου να γίνει κατανοητό το εύρος εφαρμογής και χρήσης των πληροφοριών αυτών, απαιτούνται περαιτέρω έρευνες σχετικά με τη σωστή ανάλυση, ερμηνεία και παρουσίαση των δεδομένων, τη σχέση μεταξύ των πληροφοριών που καταγράφει το σύστημα και του κινδύνου ατυχημάτων και τη δυνατότητα χρήσης αυτών των δεδομένων για τη βελτίωση της μοντελοποίησης της οδηγικής συμπεριφοράς.

2.2.3 Έξυπνα κινητά για την παρακολούθηση της οδηγικής συμπεριφοράς

Στις μέρες μας, τα έξυπνα κινητά είναι οικονομικά, αποδοτικά, εύκολα διαθέσιμα στην αγορά και μέρος της καθημερινότητας. Αυτές οι συσκευές χειρός είναι εξοπλισμένες με μια ποικιλία αισθητήρων μέσω των οποίων

μπορούν να αναπτυχθούν διάφορες εφαρμογές που σχετίζονται με **ITS (IntelligentTransportSystems),συνδεδεμένα συστήματα στόλου, καθώς και συστήματα ασφάλισης βασισμένων στη χρήση (UBI – UsageBasedInsurance)** (Eren et al. 2012, Handel et al. 2014, Husnjak et al. 2015, Johnson and Trivedi, 2011, Meseguer et al. 2013, Tselentis et al. 2016).

Το κύριο πλεονέκτημα της χρήσης έξυπνων κινητών τηλεφώνων είναι ότι **μπορούν να αποτελέσουν ένα μη παρεμβατικό περιβάλλον για τη συνεχή συλλογή πλούσιων και λεπτομερέστατων δεδομένων** σχετικά με το πραγματικό έργο της οδήγησης. Σε αντίθεση με τα τρέχοντα πειράματα οδήγησης σε πραγματικές συνθήκες, οι αισθητήρες έξυπνων κινητών τηλεφώνων είναι μια πολύ πιο βιώσιμη λύση, σε σύγκριση με στα ενσωματωμένα συστήματα παρακολούθησης οδήγησης. Η εγκατάσταση τέτοιων οργάνων είναι αρκετά δαπανηρή και δύσκολο να εγκατασταθεί και να συντηρηθεί, ειδικά όταν τα πειράματα περιλαμβάνουν μεγάλο αριθμό οχημάτων (Vlahogianni et al., 2013, 2014).Επιπλέον, τα έξυπνα κινητά τηλέφωνα είναι αρκετά διαφανείς μηχανισμοί συλλογής δεδομένων, με την έννοια ότι οι χρήστες μπορούν να στραφούν στα δεδομένα τους για να ελέγξουν τη συμπεριφορά τους ή ακόμα και να αλληλεπιδράσουν μαζί τους μέσω του κατάλληλου λογισμικού, που συνήθως αναπτύσσεται για να λειτουργεί σε ένα smartphone.

Τα έξυπνα κινητά τηλέφωνα είναι εξοπλισμένα με διάφορους αισθητήρες, όπως αισθητήρες κίνησης (π.χ. επιταχυνσιόμετρο και γυροσκόπιο), αισθητήρες θέσης (π.χ. μαγνητόμετρο), δέκτες παγκόσμιου δορυφορικού συστήματος πλοήγησης (GNSS), περιβαλλοντικοί αισθητήρες (βαρόμετρα, φωτομετρικά και θερμομέτρα), μικρόφωνο, κάμερες, κλπ. Αν και θεωρητικά, κάθε αισθητήρας μπορεί να προσθέσει γνώση στη συμπεριφορά και τις συνθήκες του οδηγού κατά τη διάρκεια της οδήγησης, η βιβλιογραφία έχει συστηματικά αντιμετωπίσει το πρόβλημα της ανίχνευσης σοβαρών συμβάντων οδήγησης χρησιμοποιώντας κυρίως το GNSS, καθώς και δεδομένα από το επιταχυνσιόμετρο, το γυροσκόπιο και το μαγνητόμετρο τριών αξόνων (Wahlstrom et al., 2015).

Με βάση τους παραπάνω αισθητήρες μπορεί να εντοπιστούν διάφορα μοτίβα οδήγησης. Μια συνοπτική ανασκόπηση των δεικτών και μετρήσεων που εξάγονται από τα έξυπνα κινητά τηλέφωνα και μπορεί να χαρακτηρίζει την οδηγική συμπεριφορά παρέχεται στο Handel et al. (2014). Τα πιο δημοφιλή από αυτά, τα οποία συχνά συναντώνται σε ασφαλιστικά προγράμματα που βασίζονται στη χρήση (UBI), περιλαμβάνουν την επιτάχυνση / πέδηση, την ταχύτητα, ως απόλυτη τιμή ή σε σχέση με ένα συγκεκριμένο όριο, δεξιά / αριστερή στροφή και στροφή (συμπεριλαμβανομένης της αλλαγής λωρίδων).

Τα καταγεγραμμένα δεδομένα μπορούν να αναλυθούν για να ανιχνεύσουν την οδηγική συμπεριφορά του οδηγού και μπορούν να ταξινομήσουν σε επιθετικές και μη επιθετικές τις διαφορετικές συμπεριφορές οδήγησης (Singh et al., 2017).

Υπάρχουν ορισμένες εφαρμογές που μπορούν να εγκατασταθούν στα έξυπνα κινητά τηλέφωνα για να επιτρέψουν τη χρήση των ενσωματωμένων αισθητήρων όπως είναι το επιταχυνσιόμετρο, το γυροσκόπιο, το μαγνητόμετρο, το GPS και ούτω καθεξής (Singh et al., 2017). Τα δεδομένα από αυτούς τους αισθητήρες μπορούν να χρησιμοποιηθούν για την ανίχνευση της ανθρώπινης δραστηριότητας. Ο αισθητήρας επιταχυνσιομέτρου χρησιμοποιείται ευρέως για διαμήκεις και πλευρικές κινήσεις, ενώ το GPS παρέχει τη θέση του τηλεφώνου ως γεωγραφικό πλάτος και γεωγραφικό μήκος. Η χρήση των smartphones για την καταγραφή της οδηγικής συμπεριφοράς, μπορεί να αποφέρει άλλο ένα πλεονέκτημα το οποίο απουσιάζει από τα IVDR, και το οποίο αφορά στην παροχή δεδομένων που αφορούν την απόσπαση της προσοχής του οδηγού λόγω χρήσης του κινητού τηλεφώνου κατά τη διάρκεια της οδήγησης (Tselentis et al., 2017).

Έτσι, τα υπάρχοντα συστήματα θα μπορούν να ανιχνεύσουν τους ελιγμούς (π.χ. επιτάχυνση / πέδηση, την ταχύτητα, ως απόλυτη τιμή ή σε σχέση με ένα συγκεκριμένο όριο, δεξιά / αριστερή στροφή (συμπεριλαμβανομένης της αλλαγής λωρίδων), καθώς και την απόσπαση της προσοχής, για να ταξινομήσουν την οδηγική συμπεριφορά του οδηγού.

Αξιολόγηση της Οδηγικής Συμπεριφοράς

Η **συχνότητα** των παραπάνω μετρήσεων σε σχέση με την **απόσταση που διανύθηκε**, την **ώρα** της ημέρας κ.λπ. μπορεί να χρησιμοποιηθεί για την **ποσοτικοποίηση της επικίνδυνης οδήγησης**, την **αξιολόγηση της συμπεριφοράς του οδηγού**, τη **διερεύνηση της επιθετικότητας** κατά την οδήγηση κ.ο.κ. (Chakravarty et al., 2013; Johnson and Trivedi, 2011; Musicant et al., 2014). Ο κανονικός τρόπος οδήγησης μπορεί να ταξινομηθεί σε διάφορες κατηγορίες, όπως κανονική ή επιθετική, ασφαλής ή επικίνδυνη, εξειδικευμένη ή ανειδίκευτη κλπ. Μπορούν να επίσης να δοθούν συστάσεις στον οδηγό για να βελτιώσει την οδήγησή του αναλόγως (feedback). Επιπροσθέτως, καθώς η οδηγική συμπεριφορά ποικίλλει από άτομο σε άτομο, και επίσης ποικίλλει λόγω ορισμένων πρόσθετων παραγόντων όπως η κατανάλωση αλκοόλ, ναρκωτικών, περιπτώσεις έκτακτης ανάγκης κλπ., υπάρχει η δυνατότητα σε τέτοιες περιπτώσεις, το αυτοματοποιημένο σύστημα παρακολούθησης του οδηγού να εμφανίζει προειδοποίηση στον οδηγό σχετικά με την οδηγική συμπεριφορά του και με τη συγκατάθεσή του, η θέση τους μπορεί να κοινοποιηθεί στις υπηρεσίες προστασίας των πολιτών.

Προφανώς, τα έξυπνα κινητά τηλέφωνα συνεπάγονται αρκετά μειονεκτήματα, όπως η αποστράγγιση της μπαταρίας, η αυθαίρετη θέση στο όχημα, τα θορυβώδη δεδομένα που περιλαμβάνουν διάφορες εκδρομές με διάφορα μέσα μεταφοράς και ούτω καθεξής. Αυτές οι αδυναμίες ενδέχεται να επηρεάσουν σημαντικά την ακρίβεια και την αξιοπιστία των συστημάτων συλλογής δεδομένων που βασίζονται σε έξυπνα κινητά τηλέφωνα. Επιπλέον, τα δεδομένα πρέπει πρώτα να φιλτραριστούν για να αποκλείσουν τον **θόρυβο που προέρχεται από ταξίδια με άλλα μέσα μεταφοράς** (λεωφορεία, μετρό κ.λπ.), **το περπάτημα**, καθώς και τα ταξίδια όπου ο ιδιοκτήτης του smartphone είναι επιβάτης (Handel et al., 2014).

Παρόλο που η ανίχνευση ταξιδιών στο αυτοκίνητο είναι μια σχετικά απλή εργασία και βασίζεται κυρίως σε πληροφορίες ταχύτητας από δεδομένα GNSS και επιταχυνσιόμετρο ή μόνο δεδομένα επιταχυνσιόμετρο, **η ανίχνευση ταξιδιού επιβατών / οδηγών είναι πολύ πιο πολύπλοκη και**

εξαιρετικά σημαντική για την ακρίβεια προσδιορίζοντας το προφίλ του οδηγού. Η ανίχνευση ταξιδιού επιβατών / οδηγών απαιτεί ακριβή δεδομένα υψηλής ανάλυσης (πιθανώς περισσότερο από 1 Hz) προκειμένου να αναλυθούν οι μικροδραστικές δραστηριότητες των χρηστών και να προσδιοριστούν τα πρότυπα που μπορούν να βοηθήσουν στην ανίχνευση της οδήγησης (Wahlstrom et al., 2015). Τέλος, η αναγνώριση της χρήσης των κινητών τηλεφώνων μπορεί να δώσει μια καλύτερη εικόνα όχι μόνο για τον εντοπισμό της συμπεριφοράς του οδηγού και την απόσπαση της προσοχής του, αλλά και για τον καθαρισμό του θορύβου που προέρχεται από τις χονδροειδείς κινήσεις του τηλεφώνου. Συγκεκριμένα, οι αισθητήρες των έξυπνων κινητών τηλεφώνων ενδέχεται να ενεργοποιηθούν από τη χρήση κινητού τηλεφώνου κατά την οδήγηση, γεγονός που επηρεάζει τη διαδικασία ανίχνευσης συμβάντων και τις τιμές κατωφλίου των κρίσιμων συμβάντων. Η αναγνώριση αυτών των κινήσεων μπορεί να μειώσει τα ψευδώς θετικά συμβάντα και να βελτιώσει την ισχύ ανίχνευσης της προσέγγισης.

2.3 Μέθοδοι Ανάλυσης των δεδομένων οδηγικής συμπεριφοράς

Στη βιβλιογραφία, υπάρχουν πολλές ερευνητικές εργασίες οι οποίες ασχολούνται με την ανάλυση δεδομένων οδηγικής συμπεριφοράς, τα οποία έχουν συλλεχθεί με μία ή συνδυασμό περισσότερων από τις προαναφερθείσες μεθόδους καταγραφής. Στην πλειοψηφία των περιπτώσεων, ο σκοπός της ανάλυσης είναι η αξιολόγηση ή κατηγοριοποίηση των οδηγών με βάση τα χαρακτηριστικά τους, σε επιθετικούς, προσεκτικούς ή επικίνδυνους, ενώ δεν υπάρχει αρκετή έρευνα βασισμένη σε τεκμήρια που να αφορά το σωστό χρονικό διάστημα (το «χρονικό πλαίσιο δειγματοληψίας») που απαιτείται για τον προσδιορισμό των τρόπων οδήγησης ή την αξιολόγηση των βελτιώσεων που επιτυγχάνονται από τα προγράμματα παρέμβασης με την πάροδο του χρόνου (Musicant et al., 2011).

2.3.1 Μέθοδοι για την Αξιολόγηση – Κατηγοριοποίηση της Οδηγικής Συμπεριφοράς

Παραδοσιακά, οι οδηγοί διακρίνονται βάσει του επιπέδου «επιθετικότητας» (π.χ., Bonsall et al., 2005). Ένας επιθετικός οδηγός παίρνει μια συμπεριφορά οδήγησης που χαρακτηρίζεται από υψηλή ταχύτητα μαζί με πολλές και ξαφνικές αλλαγές της στιγμιαίας ταχύτητας, οι οποίες προκαλούν ξαφνικές επιταχύνσεις και επιβραδύνσεις. Όσο μεγαλύτερη είναι η διακύμανση της ταχύτητας, τόσο μεγαλύτερες θα είναι οι αλληλεπιδράσεις μεταξύ των οχημάτων (Mehtar et al., 2013). Κατά συνέπεια, μπορεί να ειπωθεί ότι ένας επιθετικός οδηγός έχει μια «επικίνδυνη οδηγική συμπεριφορά». Αντίθετα, ορισμένοι ερευνητές έχουν διακρίνει μεταξύ δύο κύριων κατηγοριών οδηγών: «επιθετικοί» ή «προσεκτικοί» (π.χ., Wang et al., 2014). Ο προσεκτικός οδηγός συχνά περιγράφεται επίσης ως «μετριοπαθής» (π.χ., Taubman-Ben-Ari et al., 2004). Ένας οδηγός μπορεί να θεωρηθεί προσεκτικός όταν προσπαθεί να διατηρήσει μια σταθερή μέτρια ταχύτητα, αποφεύγοντας την απότομη επιτάχυνση και πέδηση, υποδεικνύοντας ότι έχει «ασφαλή οδηγική συμπεριφορά».

Στην πλειοψηφία τους, οι μέθοδοι που χρησιμοποιούνται εξετάζουν τις εξής παραμέτρους οδηγικής συμπεριφοράς: **συμβάντα απότομης επιτάχυνσης, συμβάντα απότομης πέδησης, οδήγηση πάνω από το όριο ταχύτητας και συμβάντα απότομης αλλαγής λωρίδας/κατεύθυνσης** (Toledo et al., 2007, Musicant et al., 2011, Tselentis et al., 2017, Shichrur et al., 2014, κ.α.).

Τέτοια γεγονότα θεωρούνται συνήθως οι παραβιάσεις κατωφλίων (thresholds) που επιβάλλονται στην επιτάχυνση του οχήματος. Οι διάφορες συνεισφορές στη βιβλιογραφία χρησιμοποιούν διαφορετικά όρια. Ενδεικτικά, ο Paefgen et al. (2012) χρησιμοποιεί διαμήκη κατώτατα όρια $a_x < 0,1g$ για απότομη πέδηση, $a_x > 0,1g$ για απότομη επιτάχυνση και πλευρική επιτάχυνση $|a_y| > 0,2g$ για απότομη αλλαγή κατεύθυνσης, ενώ οι Baldwin, Duncan, and West (2004) χρησιμοποιούν ως όριο **0,15g σε όλες τις κατευθύνσεις**.

Συνεπώς, ένας οδηγός κατατάσσεται σε μία από τις παραπάνω κατηγορίες, χρησιμοποιώντας τα παραπάνω όρια για την ανίχνευση των επικίνδυνων συμπεριφορών κατά τη διάρκεια της οδήγησης. Η κατάταξη αυτή όμως γίνεται χρησιμοποιώντας διαφορετικές (στατιστικές) μεθόδους, όπως είναι τα διαγράμματα “g-g”, η ανάπτυξη προφίλ συμπεριφοράς (DriverBehaviorProfiles – DBP), η ομαδοποίηση k-μέσων (k-means clustering), η χρήση του μέτρου μεταβλητότητας συμπεριφοράς (volatility) κ.α. Παρακάτω, αναφέρονται επιγραμματικά κάποιες από τις μεθόδους αυτές.

Μεταξύ όλων των διαθέσιμων μεθόδων της επιστημονικής βιβλιογραφίας, ένας τρόπος για τον χαρακτηρισμό της συμπεριφοράς του οδηγού είναι το διάγραμμα “g-g”, που δείχνει τις διαμήκεις και πλευρικές επιταχύνσεις στους άξονες y και x, κανονικοποιημένες σε σχέση με τη βαρύτητα, που καταγράφονται σε ένα όχημα κατά τη διάρκεια μιας πραγματικής δοκιμής στο δρόμο (Hisaoaka et al., 1999). Όπως τονίστηκε από τους Da Lio et al. (2005), αυτό το διάγραμμα αντιπροσωπεύει ένα πολύ αποτελεσματικό τρόπο για τον χαρακτηρισμό της απόδοσης/συμπεριφοράς του οδηγού ενός οχήματος, διότι οι τιμές διαμήκους και πλευρικής επιτάχυνσης μπορούν να θεωρηθούν ως η σύνθεση μεταξύ της δυναμικής του οχήματος και της συμπεριφοράς οδήγησης που επιλέγεται, σύμφωνα με την αντίληψη κίνησης του οδηγού και το επίπεδο κινδύνου που αυτός αποδέχεται. Ωστόσο, είναι απαραίτητο να ληφθεί υπόψη ότι η ταχύτητα είναι επίσης σημαντική για να χαρακτηρίσει τη συμπεριφορά των οδηγών και ότι η ταχύτητα και οι επιταχύνσεις αλληλοσυνδέονται.

Μία ακόμα μέθοδος για τον χαρακτηρισμό της συμπεριφοράς του οδηγού είναι τα προφίλ οδηγικής συμπεριφοράς (Driver Behavior Profiles–DBP). Ένα DBP είναι ένα σύνθετο κανονικοποιημένο μέτρο / βαθμολογία της συμπεριφοράς του οδηγού που χρησιμεύει ως υποκατάστατο για την εκτίμηση του κινδύνου σύγκρουσης, καθώς συνδυάζει διάφορα μέτρα κινδύνου σε μια κοινή κλίμακα εντός και μεταξύ των οδηγών, ικανοποιώντας έτσι την ανάγκη για ένα ενιαίο μέτρο συμπεριφοράς του οδηγού σε σχέση με την ασφάλεια, ενώ παράλληλα καταγράφει τη μεταβλητότητα και την πληθώρα των πτυχών που ενσωματώνονται στο έργο οδήγησης (Ellison et al., 2015). Το μέτρο

χρησιμοποιεί εμπειρικά δεδομένα σχετικά με τη συμπεριφορά του οδηγού, όπως η ταχύτητα, η απότομη επιτάχυνση και η απότομη πέδηση, λαμβάνοντας υπόψη τις συχνότητες και τα μεγέθη με τα οποία συμβαίνουν. Αυτό επιτρέπει τη σύγκριση των οδηγών μεταξύ τους και τη σύγκριση της συμπεριφοράς του ίδιου οδηγού καθόλη τη διάρκεια του χρόνου και του χώρου.

Άλλη μια μέθοδος που χρησιμοποιείται συχνά από τους ερευνητές είναι η ομαδοποίηση k -μέσων. Η ομαδοποίηση είναι το έργο της διαίρεσης ενός συνόλου δειγμάτων σε διάφορες ομάδες (clusters) έτσι ώστε τα δείγματα εντός της ίδιας ομάδας να είναι παρόμοια. Η ομαδοποίηση k -μέσων στοχεύει στη διαίρεση n παρατηρήσεων σε k ομάδες στις οποίες κάθε παρατήρηση ανήκει στην ομάδα με την πλησιέστερη μέση τιμή, που χρησιμεύει ως αντιπροσωπευτική της ομάδας (Yeo, Smith, Willis, & Brooks, 2001). Οι Mantouka et al., (2017) χρησιμοποίησαν την ομαδοποίηση k -μέσων σε συνδυασμό με τη μεταβλητότητα συμπεριφοράς (volatility) για τον χαρακτηρισμό της συμπεριφοράς του οδηγού. Συγκεκριμένα, στην παραπάνω έρευνα εφαρμόστηκε μια ομαδοποίηση δύο σταδίων για την ανίχνευση, σε πρώτη φάση της επιθετικότητας με βάση την εμφάνιση απότομων συμβάντων (επιτάχυνση και επιβράδυνση), ενώ στη συνέχεια, εντοπίστηκαν επιπρόσθετες επικίνδυνες συμπεριφορές οδήγησης, ήτοι η απόσπαση της προσοχής (χρήση κινητού τηλεφώνου) και η ανάληψη κινδύνου (εκτέλεση παραβιάσεων ορίων ταχύτητας) τόσο των επιθετικών όσο και των μη επιθετικών καταστάσεων οδήγησης. Επιπλέον διεξήχθη ανάλυση των ταξιδιών κάθε οδηγού, προκειμένου να ανιχνευθεί η μέση συμπεριφορά τους, είτε βελτιώνουν είτε επιδεινώνουν τον τρόπο οδήγησης τους, καθώς και αν παρουσιάζουν σταθερή ή μεταβαλλόμενη συμπεριφορά (Volatility measure).

2.3.2 Μέθοδοι Ποσοτικοποίησης Δεδομένων

Όπως αναφέρθηκε και προηγουμένως, δεν υπάρχει αρκετή έρευνα που να αφορά το σωστό / βέλτιστο χρονικό διάστημα που απαιτείται να παρατηρηθεί

ένας οδηγός ώστε να προσδιοριστεί η συμπεριφορά οδήγησης ή η αξιολόγηση των βελτιώσεων που επιτυγχάνονται από τα προγράμματα παρέμβασης με την πάροδο του χρόνου (Musicant et al., 2011). Μια πτυχή έμμεσα σχετιζόμενη με αυτό το θέμα έχει διερευνηθεί από τη δεκαετία του 1960 (Perkins, Harris, 1968, NCHRP, 1999), με έρευνες που προσπαθούν να απαντήσουν στο ερώτημα πόσο χρόνο πρέπει να παρατηρείται ένας χώρος για να ληφθούν αξιόπιστες εκτιμήσεις για τα ποσοστά σύγκρουσης (Robertson et al., 1994, Parker and Zegeer, 1988).

Αναθεωρώντας την πρόσφατη βιβλιογραφία σχετικά με την αξιολόγηση της συμπεριφοράς οδήγησης, διαπιστώνεται ότι, όχι μόνο δεν υπάρχει συναίνεση σχετικά με τον εύλογο χρόνο δειγματοληψίας για μετρήσεις με χρήση IVDR ή έξυπνων κινητών τηλεφώνων στην αξιολόγηση των οδηγών, αλλά και μόνο λίγες έρευνες έχουν επικεντρωθεί στον απαιτούμενο χρόνο δειγματοληψίας (Shichrur et al., 2014).

Ο Shichrur et al. (2014) χρησιμοποίησε **δεδομένα από IVDR** τα οποία ήταν εγκατεστημένα στα οχήματα 64 επαγγελματιών οδηγών ταξί για τον καθορισμό του χρονικού πλαισίου δειγματοληψίας που απαιτείται για την επίτευξη ενός σταθερού μοτίβου οδήγησης χρησιμοποιώντας το μέσο ποσοστό συμβάντων, αναλύοντας τα πρότυπα οδήγησης και τις αλλαγές στη συμπεριφορά των οδηγών με την πάροδο του χρόνου.

Πραγματοποιήθηκαν δύο στατιστικές αναλύσεις στα δεδομένα από τα IVDR. Στην πρώτη ανάλυση χρησιμοποιήθηκε ένας **δείκτης σύγκλισης** για να καθορίσει σε ποιο βαθμό τα δεδομένα παρουσιάζουν συγκλίνοντα ρυθμό συμβάντων, προκειμένου να προσδιοριστεί ο χρόνος στον οποίο σταθεροποιείται ο ρυθμός των ανεπιθύμητων συμβάντων οδήγησης. Κατά την δεύτερη ανάλυση, η οποία αφορούσε στο **αναμενόμενο σφάλμα**, εξετάστηκε το μέγεθος του αναμενόμενου σφάλματος το οποίο παράγεται από ένα δεδομένο χρονικό πλαίσιο δειγματοληψίας όταν αυτό χρησιμοποιείται για την αξιολόγηση του μοτίβου οδήγησης.

Η χρήση του δείκτη σύγκλισης έδειξε ότι η παρακολούθηση των οδηγών για λιγότερο από 100 ώρες οδήγησης δεν παρέχει αξιόπιστη μέτρηση. Καθώς όμως αυτό δεν απαντούσε ακριβώς το αρχικό ερώτημα των ερευνητών για τον καθορισμό του χρονικού πλαισίου παρατήρησης των οδηγών, χρησιμοποιήθηκε και μια άλλη μέθοδος υπολογισμού, αυτή του μέσου αναμενόμενου σφάλματος που δίνεται στην «περικοπή» του δείγματος χρόνου κάθε 100 ώρες οδήγησης. Η ανάλυση αυτή έδειξε ότι η **συλλογή δείγματος περίπου 300 ωρών** ανά οδηγό αποτελεί σχετικά σταθερό και **αξιόπιστο μέτρο** για την αξιολόγηση του μέσου ποσοστού συμβάντος του οδηγού (αναμενόμενο σφάλμα = 10%). Αντιθέτως, η δειγματοληψία με **λιγότερες από 100 ώρες** οδήγησης **δεν οδηγεί σε αξιόπιστη μέτρηση** για την εκτίμηση του ποσοστού συμβάντος του οδηγού (αναμενόμενο σφάλμα = 25%), ενώ η δειγματοληψία **μεταξύ 100 και 300 ωρών** μπορεί να οδηγήσει σε αξιόπιστα αποτελέσματα, αλλά **συνιστάται λιγότερο**.

2.4 Γενικά Συμπεράσματα Βιβλιογραφικής Ανασκόπησης

Η κατανόηση, η καταγραφή και η ανάλυση της οδηγικής συμπεριφοράς έχει αποτελέσει αντικείμενο πλήθους επιστημονικών ερευνών, καθώς επηρεάζει όχι μόνο την οδική ασφάλεια, αλλά και την οικονομία και την ευημερία της κοινωνίας συνολικά. Πολλές διαφορετικές μέθοδοι έχουν παρουσιαστεί, τόσο για την καταγραφή όσο και για την αξιολόγησή της από τους ερευνητές, με τις πλέον πρόσφατες και «επαναστατικές» να συμπεριλαμβάνουν την αξιοποίηση της τεχνολογίας των έξυπνων κινητών τηλεφώνων, σαν μια νέα, βιώσιμη εναλλακτική των παλαιότερων μεθόδων καταγραφής. Τα πρώτα αποτελέσματα από σχετικές εφαρμογές (Tselentis et al., 2017, Theofilatos et al., 2017, Araújo et al., 2012, Vlahogianni and Barmounakis, 2017) επιβεβαίωσαν την αποτελεσματικότητα χρησιμότητας αυτών των μεγάλων συστημάτων συλλογής δεδομένων.

Παρ' όλα αυτά, δεν έχει προσδιορισθεί ακόμη ο απαιτούμενος χρόνος παρατήρησης κάθε οδηγού για την εκτίμηση της οδηγικής συμπεριφοράς του.

Λαμβάνοντας υπόψιν ότι, αφενός ένα μικρό δείγμα δεδομένων ενέχει τον κίνδυνο να οδηγήσει την έρευνα σε αμφισβητήσιμα αποτελέσματα, και αφετέρου, ένα μεγάλο δείγμα δεδομένων απαιτεί υψηλό υπολογιστικό κόστος ανάλυσης, προκύπτει η σημασία της παραπάνω παρατήρησης.

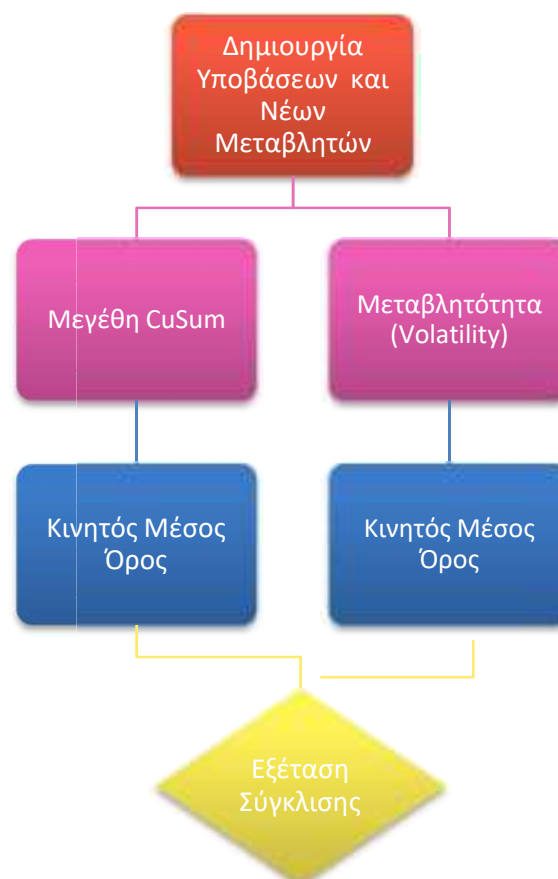
Κατά συνέπεια, θα ήταν πολύτιμο να καθοριστεί το απαιτούμενο χρονικό διάστημα καταγραφής δεδομένων οδήγησης, προκειμένου να αποκτηθεί μια ολοκληρωμένη εικόνα για κάθε οδηγό, όπου ο ρυθμός των μετρήσεων που περιγράφονται παραπάνω (απότομες επιταχύνσεις/επιβραδύνσεις, υπέρβαση του ορίου ταχύτητας, απόσπαση προσοχής λόγω χρήσης κινητού τηλεφώνου) ανά χιλιόμετρο (ή ανά ώρα) συγκλίνει σε μια σταθερή τιμή.

3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

3.1 Ροή Εργασιών

Στο εδάφιο αυτό αναπτύσσεται αναλυτικά η μεθοδολογία που ακολουθήθηκε για την εξαγωγή των τελικών αποτελεσμάτων της παρούσας Διπλωματικής Εργασίας. Χρησιμοποιήθηκαν διάφορες μέθοδοι και στατιστικά εργαλεία, τόσο για την αρχική επεξεργασία των δεδομένων, όσο και για την εξαγωγή των συμπερασμάτων. Στη συνέχεια επεξηγούνται αναλυτικά οι λόγοι που οδήγησαν στην επιλογή της κάθε μεθόδου καθώς επίσης και ο τρόπος με τον οποίο χρησιμοποιήθηκαν στην εργασία αυτή.

Στο Διάγραμμα 3 παρατίθεται σχηματική απεικόνιση της προτεινόμενης μεθοδολογικής προσέγγισης για την ανάδειξη ενός χρονικού σημείου ως το σημείο σύγκλισης της συμπεριφοράς ενός οδηγού.



Διάγραμμα 3.1: Σχηματική Απεικόνιση Μεθοδολογικής Προσέγγισης

3.1.1 Προκαταρκτική Επεξεργασία Δεδομένων

Η γενική «μορφή» της βάσης δεδομένων, η οποία αποτελείται από **ξεχωριστά ταξίδια-διαδρομές**, απαιτεί την δημιουργία ενός **τύπου «ομαδοποίησης» των δεδομένων** προκειμένου να γίνει εφικτή η παρακολούθηση της **μεταβολή** των μεγεθών των χαρακτηριστικών οδηγικής συμπεριφοράς με την πάροδο του χρόνου. Για τον λόγο αυτό, επομένως, ακολουθήθηκε αρχικά η μεθοδολογία **διαχωρισμού των συνολικών ταξιδιών ανάλογα με την διάρκειά τους**.

Χρησιμοποιώντας το πακέτο ελεύθερου λογισμικού **«RStudio»**, συντάχθηκε κώδικας ο οποίος δημιουργούσε **«υποβάσεις»** της αρχικής βάσης δεδομένων, **με βάση την διάρκεια των ταξιδιών**. Για παράδειγμα, η υποβάση «DUR5» περιέχει όλα τα ταξίδια όλων των οδηγών τα οποία όμως έχουν συνολική καθαρή διάρκεια οδήγησης μικρότερη των 5 λεπτών, η «DUR10» περιέχει όλα εκείνα τα ταξίδια των οποίων η διάρκεια ήταν μεταξύ 5 και 15 λεπτών (μέση τιμή διάρκειας υποβάσης: 10 λεπτά) και η «DUR20» αντίστοιχα τα ταξίδια των οποίων η διάρκεια ήταν μεταξύ 15 και 25 λεπτών (μέση τιμή διάρκειας υποβάσης: 20 λεπτά). Ο χωρισμός αυτός πραγματοποιήθηκε για τις παραπάνω διάρκειες καθότι ο αριθμός των ταξιδιών με διάρκεια μεγαλύτερη των 25 λεπτών ήταν σημαντικά μικρότερος, με αποτέλεσμα αρκετοί οδηγοί να έχουν λιγότερα από 30 ταξίδια στις υποβάσεις αυτές, ή και πολλές φορές κανένα απολύτως. Συνεπώς, μια τέτοια διαδικασία ήταν απαραίτητη, ώστε τα δεδομένα να έχουν τα **χαρακτηριστικά των χρονοσειρών** και να μπορούν να μελετηθούν ως τέτοια.

Στη συνέχεια οι παραπάνω υποβάσεις χωρίστηκαν επιπλέον και **ανά οδηγό**, ώστε για παράδειγμα η υποβάση με την ονομασία: «DUR10_USER_4» να περιέχει μόνο τα ταξίδια του οδηγού με κωδική ονομασία «4», τα οποία είχαν μέση καθαρή διάρκεια 10 λεπτά. Με τον τρόπο αυτό η αρχική βάση δεδομένων κατέστη δυνατό να χρησιμοποιηθεί για τους σκοπούς της παρούσας Διπλωματικής Εργασίας και εν συνεχεία να υποστεί την απαιτούμενη επεξεργασία για την εξαγωγή των αποτελεσμάτων.

3.1.2 Σωρευτικά Αθροίσματα Οδηγικών Παραμέτρων

Το επόμενο βήμα που ακολουθήθηκε αφορούσε στη δημιουργία νέων μεταβλητών των μεγεθών που περιγράφουν την οδηγική συμπεριφορά, και συγκεκριμένα των **σωρευτικών αθροισμάτων τους (Cumulative Sums)**. Τα σωρευτικά αθροίσματα (ή για συντομία CuSum) χρησιμοποιήθηκαν για να γίνει πιο εμφανής η **τάση** των οδηγικών χαρακτηριστικών **να σταθεροποιούνται ή να κυμαίνονται γύρω από μία σταθερή τιμή με την πάροδο του χρόνου**. Η τάση αυτή είναι πολύ πιο εύκολα αντιληπτή με την οπτικοποίηση των μεταβλητών CuSum σε διαγράμματα στα οποία ο οριζόντιος άξονας παριστάνει τον χρόνο ή τον αριθμό ταξιδιών ή τα διανυθέντα χιλιόμετρα μέχρι την στιγμή εκείνη.

Μεταβλητές CuSum δημιουργήθηκαν τόσο για τα μεγέθη των χαρακτηριστικών οδηγικής συμπεριφοράς, όσο και για τα μεγέθη της συνολικής καθαρής διάρκειας και συνολικής καθαρής απόστασης οδήγησης. Αυτό έγινε, διότι στην τελική ανάλυση χρησιμοποιήθηκαν **παράγωγα μεγέθη ανηγμένα στην συνολική διάρκεια ή απόσταση οδήγησης**, ώστε τα **αποτελέσματα** μεταξύ των διαφορετικών οδηγών να είναι **συγκρίσιμα**.

Εκτός από την παραπάνω μέθοδο, χρησιμοποιήθηκαν επίσης και τα απλά στατιστικά **όρια**:

$$\text{Μέση Τιμή} \pm 1.0 * \text{Τυπική Απόκλιση} \quad (1)$$

των προαναφερθέντων μεταβλητών, ώστε να ελεγχθεί αν και πότε η χρονοσειρά της εκάστοτε μεταβλητής είναι εντός των ορίων αυτών.

3.1.3 Μέθοδος Μεταβλητότητας Συμπεριφοράς

Ένα ακόμα μέγεθος το οποίο χρησιμοποιήθηκε για την αξιολόγηση της συγκλίνουσας ή μη συγκλίνουσας συμπεριφοράς των οδηγών ήταν το

μέγεθος της **μεταβλητότητας της συμπεριφοράς (volatility)**. Το στατιστικό αυτό μέτρο εφαρμόστηκε σαν «φίλτρο» πάνω στις νέες **παράγωγες μεταβλητές** που δημιουργήθηκαν και αναφέρονται στο προηγούμενο εδάφιο. Σκοπός ήταν η **διερεύνηση της μεταβλητότητας** των εν λόγω μεγεθών, κυρίως κατά την φάση στην οποία φαίνεται να έχει αρχίσει η σύγκλιση της συμπεριφοράς τους.

Για την εφαρμογή του στατιστικού αυτού μέτρου χρησιμοποιήθηκε ένα **«κινούμενο παράθυρο»** το οποίο ήταν σταθερό και ίσο με **20 ταξίδια**. Ο αριθμός αυτός ορίστηκε κάνοντας την εύλογη **παραδοχή** ότι κανένας οδηγός και κανένα χαρακτηριστικό δεν συγκλίνει νωρίτερα από τα 20 ταξίδια παρατήρησης, και άρα χρειάζεται να παρατηρήσουμε έναν οδηγό σίγουρα για περισσότερα από 20 ταξίδια, πράγμα το οποίο επαληθεύεται και αργότερα από τα αποτελέσματα της παρούσας Εργασίας.

Καθώς το μέγεθος του volatility είναι **καθαρός αριθμός** ο οποίος δείχνει μεταβλητότητα, εκείνο το οποίο εξετάζεται είναι κατά πόσο αυτός ο αριθμός είναι «μικρός» ή «μεγάλος». **Μεγάλη τιμή volatility** σημαίνει ότι υπάρχει **μεγάλη μεταβλητότητα** του εκάστοτε υπό μελέτη χαρακτηριστικού, ενώ το αντίθετο σημαίνει η μικρή τιμή. Παρατηρείται επομένως μια **τάση** του μεγέθους του volatility **να σταθεροποιείται ή να κυμαίνεται γύρω από μια σταθερή τιμή όσο αυξάνεται ο χρόνος παρατήρησης των οδηγών**. Η τάση αυτή είναι πολύ πιο εύκολα αντιληπτή με την οπτικοποίηση του μεγέθους του volatility σε διάγραμμα στο οποίο ο οριζόντιος άξονας παριστάνει τον αριθμό ταξιδιών μέχρι τη στιγμή εκείνη.

Και σε αυτήν την περίπτωση χρησιμοποιήθηκαν επιπλέον και τα απλά στατιστικά **όρια**:

$$\text{Μέση Τιμή} \pm 1.0 * \text{Τυπική Απόκλιση} \quad (2)$$

αυτή τη φορά του μεγέθους του volatility, ώστε να ελεγχθεί αν και πότε η χρονοσειρά των τιμών του είναι εντός των ορίων αυτών.

3.1.4 Κινητός Μέσος Όρος (MovingAverage)

Καθώς τα δεδομένα μας έχουν τα χαρακτηριστικά των χρονοσειρών, ένας πολύ καλός τρόπος να μελετηθεί η **μεταβλητότητά τους σε συνάρτηση με τον χρόνο**, είναι η χρήση του **Κινητού Μέσου Όρου (MovingAverage)**. Ο Κινητός Μέσος Όρος (MA) χρησιμοποιήθηκε τόσο στις νέες μεταβλητές **CuSum**, όσο και στο μέγεθος του **Volatility**, χρησιμοποιώντας και εδώ το χρονικό παράθυρο των 20 ταξιδιών.

Η χρήση του Κινητού Μέσου Όρου προσφέρει την δυνατότητα της **«ομαλοποίησης» των δεδομένων** (και άρα και των διαγραμμάτων), καθώς όσο πιο μεγάλο είναι το χρονικό παράθυρο, τόσο λιγότερο επηρεάζεται από τις ακραίες τιμές των μεγεθών που εξετάζονται (τοπικά μέγιστα και ελάχιστα), και επομένως «φιλτράρονται» ακόμα περισσότερο τα δεδομένα. Συνεπώς, κατ' αυτόν τον τρόπο προσεγγίζεται ακόμα καλύτερα η μέση συμπεριφορά στην οποία συγκλίνει ο κάθε οδηγός.

Ποσοστιαία Μεταβολή του Κινητού Μέσου Όρου (% ChangeMA)

Παρόλα αυτά, όπως είναι φυσικό, συνεχίζουν να υπάρχουν διαφορές ανάμεσα στις τιμές του Κινητού Μέσου Όρου για διαδοχικά ταξίδια του ίδιου οδηγού. Ένα ακόμα μέτρο το οποίο χρησιμοποιήθηκε ώστε να γίνει δεκτό ότι από ένα συγκεκριμένο σημείο και μετά ο οδηγός παρουσιάζει συγκλίνουσα συμπεριφορά είναι η **Ποσοστιαία Μεταβολή του Κινητού Μέσου Όρου (% ChangeMA)** από ταξίδι σε ταξίδι.

Για τον καθορισμό του παραπάνω χρονικού σημείου, απαιτείται τουλάχιστον για **5 διαδοχικά ταξίδια** του ίδιου οδηγού, η μέγιστη ποσοστιαία μεταβολή του MA να είναι ίση με **1,50%**. Η παραπάνω παραδοχή προέκυψε ως ελάχιστο αποδεκτό όριο από την εμπειρία και ύστερα από δοκιμές στα δεδομένα, και επ' αυτής απαιτείται περαιτέρω διερεύνηση.

Τοπικά Ακρότατα

Για να ταυτοποιηθεί ένα χρονικό σημείο ως το ζητούμενο σημείο από το οποίο και μετά ο οδηγός διατηρεί μια περίπου σταθερή συμπεριφορά, θα πρέπει το σημείο αυτό να αποτελεί επίσης και **τοπικό ακρότατο** (μέγιστο ή ελάχιστο) **της χρονοσειράς του MA**. Το κριτήριο αυτό εξασφαλίζει ότι οι γειτονικές τιμές του MA είναι μικρότερες ή μεγαλύτερες του σημείου που έχει επιλεγεί, και επομένως δεν ανήκει σε κάποια ακολουθία σημείων τα οποία έχουν μια συγκεκριμένη τάση (π.χ. αύξουσα ή φθίνουσα).

3.2 Θεωρητικό Υπόβαθρο

Σε αυτό το εδάφιο παρουσιάζεται το θεωρητικό υπόβαθρο στο οποίο βασίστηκε η ανάλυση του αντικειμένου της Διπλωματικής Εργασίας καθώς και βασικές στατιστικές μέθοδοι ανάλυσης. Για τον σκοπό της παρούσας εργασίας εφαρμόστηκαν οι μέθοδοι Σωρευτικών Αθροισμάτων (CuSum), Μεταβλητότητας Συμπεριφοράς (Volatility) και χρησιμοποιήθηκαν εργαλεία όπως ο Κινητός Μέσος Όρος (MovingAverage) και οι Ποσοστιαίες Μεταβολές, ώστε να επιτευχθούν τα τελικά αποτελέσματα. Μεγάλο τμήμα των θεωρητικών στοιχείων αυτών έχει αναπτυχθεί και σε προηγούμενες διπλωματικές εργασίες.

3.2.1 Βασικές Έννοιες Χρονοσειρών

Η **χρονοσειρά** ορίζεται ως το **σύνολο διαδοχικών τιμών ενός χαρακτηριστικού** (παρατηρήσεων), οι οποίες **καταγράφονται με τη σειρά που παράγονται** για ένα χρονικό διάστημα. Πιο συγκεκριμένα, δοθέντος ενός χαρακτηριστικού A , μια χρονοσειρά είναι το σύνολο n παρατηρηθέντων ζευγών:

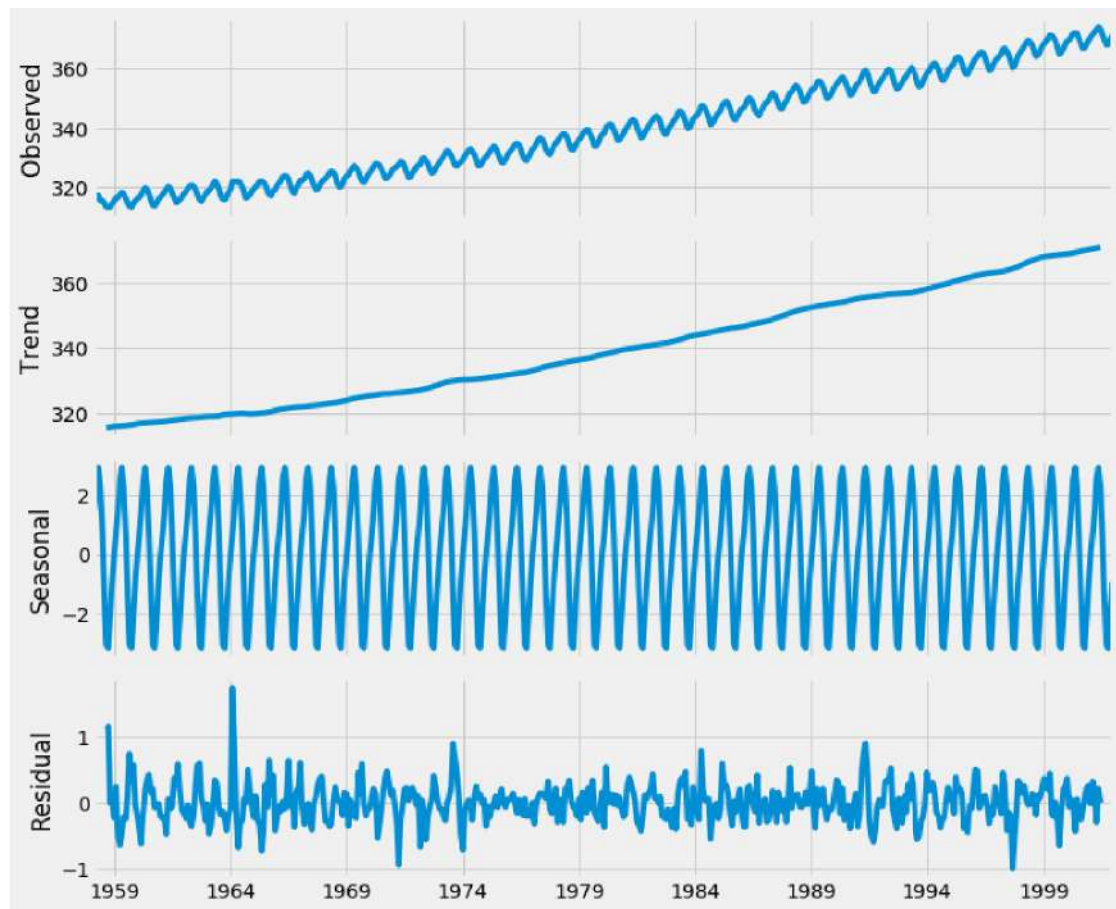
$$\{(t_1, a_1), (t_2, a_2), \dots, (t_n, a_n)\} \quad (3)$$

όπου

$$T = \{1, 2, \dots, n\} \quad (4)$$

είναι μια **αλληλουχία διαδοχικών χρονικών στιγμών** (π.χ. ημέρα, μήνας κτλ.), ενώ **ανείναι μια παρατηρηθείσα τιμή** επί του χαρακτηριστικού A για τη χρονική στιγμή n ως παρατήρηση στο χρόνο αυτό (Aggarwal 2015, Dunham 2003, Han & Kamber 2001). Τελευταίο στοιχείο της μεταβλητής λοιπόν είναι αυτό που καταγράφηκε τελευταίο. Για παράδειγμα, τελευταίο στοιχείο της μεταβλητής που καταγράφει τα απότομα φρεναρίσματα διαδοχικών ταξιδιών

μέσης διάρκειας 10 λεπτών, είναι ο αριθμός απότομων φρεναρισμάτων του τελευταίου καταγεγραμμένου ταξιδιού.



Εικόνα 3.1: Παράδειγμα χρονοσειράς, όπου παρατηρείται και η τάση, η εποχικότητα και τα κατάλοιπα (Πηγή:DigitalOcean)

Οι θεμελιώδεις έννοιες στην ανάλυση χρονοσειρών είναι δύο: η στασιμότητα και η αυτοσυσχέτιση.

Στασιμότητα

Η στασιμότητα αναφέρεται στην **τάση** και την **μεταβλητότητα** μιας χρονοσειράς. Τα πρότυπα χρονοσειρών εφαρμόζονται σε οριζόντια, ή αλλιώς με στατιστικούς όρους σε στάσιμα δεδομένα (Karlaftis, Mannering and Washington, 2010). Για να είναι μια χρονοσειρά στάσιμη, θα πρέπει τα

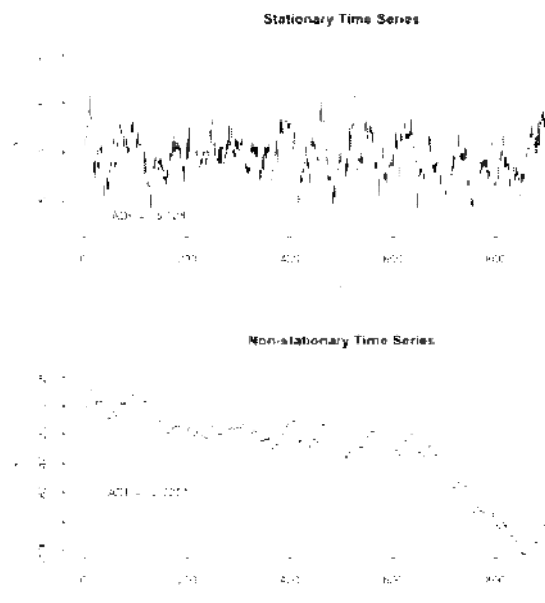
βασικότερα χαρακτηριστικά της, δηλαδή η **μέση τιμή** και η **διασπορά** της, να παραμένουν σταθερά και **να μην μεταβάλλονται με την πάροδο του χρόνου** και η συνδιακύμανση μεταξύ των τιμών της σε δύο χρονικά σημεία, **να εξαρτάται μόνο από την απόσταση τους** και όχι από τον χρόνο. Τα δεδομένα με λίγα λόγια, θα πρέπει να μην παρουσιάζουν καμία τάση ή εποχικότητα και οι τιμές που λαμβάνουν να είναι μόνο συνάρτηση της μεταξύ τους απόστασης και όχι του ίδιου του χρόνου. Με άλλα λόγια θα μπορούσε να ειπωθεί ότι για να είναι μια χρονοσειρά στάσιμη, θα πρέπει οι ιδιότητές της να μην μεταβάλλονται από μια αλλαγή στην αρχή μέτρησης του χρόνου. Μαθηματικά, τα προαναφερθέντα βασικά χαρακτηριστικά μιας χρονοσειράς, με $\{X_t, t \in T\}$ εκφράζονται ως εξής:

Συνάρτηση Μέσης Τιμής: $\mu(t) = E[X_t], t \in T$ (5)

Συνάρτηση Διασποράς: $\sigma^2(t) = V[X_t] = E[(X_t - \mu_t)^2], t \in T$ (6)

Συνάρτηση Αυτοδιακύμανσης:

$\gamma(t, h) = \text{Cov}(X_t, X_{t+h}) = E[(X_t - \mu_t)(X_{t+h} - \mu_{t+h})], t, h \in T$ (7)



Εικόνα 3.2: Σύγκριση στάσιμης (πάνω) και μη-στάσιμης (κάτω) χρονοσειράς (Πηγή:Wikipedia)

Αυτοσυσχέτιση

Η αυτοσυσχέτιση εκφράζει τον **συσχετισμό μεταξύ των παρελθοντικών και μελλοντικών τιμών** μιας χρονοσειράς. Στην ανάλυση χρονολογικών σειρών η σημασία της συνάρτησης αυτοσυσχέτισης είναι πολύ μεγάλη, γιατί δείχνει τόσο τον βαθμό, όσο και το μήκος ή τη χρονική διάρκεια της μνήμης της στοχαστικής διαδικασίας. Η **αυτοσυσχέτιση** ρ_τ για κάποια **χρονική υστέρηση** τ , μπορεί να εκφραστεί ως ο **συντελεστής συσχέτισης** δύο στοιχείων μιας χρονοσειράς που **απέχουν μεταξύ τους τ χρονικές υστερήσεις** και εκτιμάται από τη χρονοσειρά ως:

$$\rho_\tau = \text{Corr}(X_t, X_{t+\tau}) = \frac{\sum_{t=\tau+1}^n (X_t - \bar{X})(X_{t-\tau} - \bar{X})}{\sum_{t=\tau+1}^n (X_t - \bar{X})^2} \quad (8)$$

Να σημειωθεί ότι ο ορισμός της **αυτοσυσχέτισης έχει νόημα** μόνο όταν πρόκειται **για στάσιμη χρονοσειρά**. Όταν δεν πρόκειται για στάσιμη χρονοσειρά, η αυτοσυσχέτιση δε μπορεί να οριστεί ως συνάρτηση της υστέρησης, αλλά ορίζεται για κάθε χρονική υστέρηση t . Οι τιμές της αυτοσυσχέτισης βρίσκονται στο διάστημα $[-1,1]$. Όταν η αυτοσυσχέτιση παίρνει τιμές κοντά στο 0 σημαίνει ότι τα δεδομένα της χρονοσειράς είναι τελείως τυχαία και δεν σχετίζονται μεταξύ τους.

Μερική Αυτοσυσχέτιση

Μια άλλη σημαντική έννοια που χρησιμοποιείται και εφαρμόζεται στη θεωρία χρονοσειρών είναι αυτή της μερικής αυτοσυσχέτισης. Η μερική αυτοσυσχέτιση εκφράζει τη συσχέτιση μεταξύ δύο σημείων Y_t και $Y_{t-\tau}$ μιας χρονοσειράς **αφού έχει αφαιρεθεί ήδη η επιρροή των υπολοίπων σημείων** $Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_{t-\tau+1}$ σε αυτές. Η μερική αυτοσυσχέτιση Φ_π ορίζεται ως εξής:

$$\Phi_\pi = \frac{r_\tau - \sum_{t=1}^{\tau-1} \Phi_{\tau-1,t} r_{t-\tau}}{1 - \sum_{t=1}^{\tau-1} \Phi_{\tau-1,t} r_t} \quad (9)$$

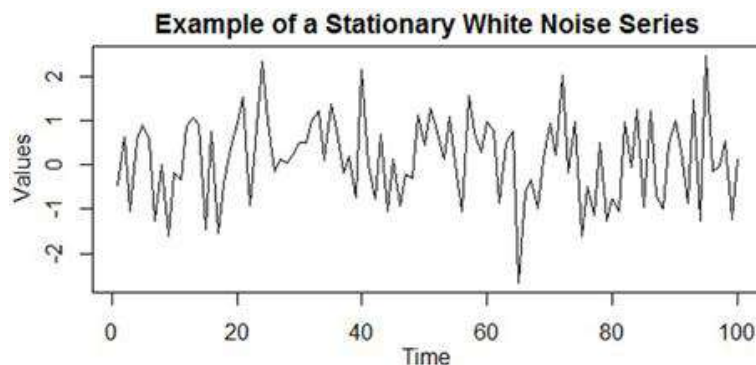
Λευκός Θόρυβος

Τέλος, μια άλλη σημαντική έννοια στη θεωρία χρονοσειρών είναι αυτή του λευκού θορύβου. Η έννοια αυτή έχει να κάνει με τα δυνατά σχήματα ή πρότυπα που μπορεί να λάβει μια μεταβλητή, όπως για παράδειγμα κύκλος, ευθεία, παραβολή κτλ. Μια σειρά ονομάζεται λευκός θόρυβος όταν δεν έχει κανένα ευκρινές σχήμα και η κατανομή της είναι τελείως τυχαία. Μαθηματικά, μια σειρά Y_t είναι λευκός θόρυβος εάν ο μέσος και η διακύμανσή της είναι αμετάβλητα και οι τιμές της δεν παρουσιάζουν αυτοσυσχέτιση.

$$\mu(t) = E[X_t] = 0 \quad (10)$$

$$\sigma^2(t) = V[X_t] = 0 \quad (11)$$

$$\gamma(t, h) = \text{Cov}(X_t, X_{t+h}) = 0 \quad (12)$$

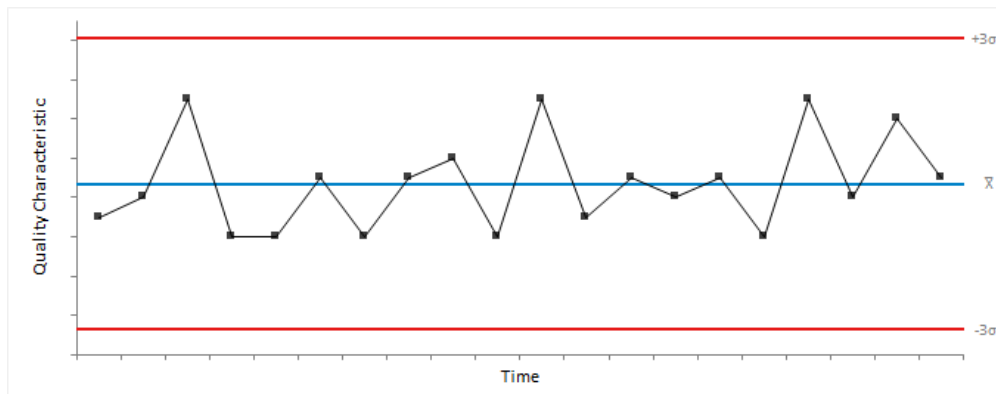


Εικόνα 3.3: Παράδειγμα Λευκού Θορύβου (Πηγή: ResearchGate)

3.2.2 Μέθοδος Σωρευτικών Αθροισμάτων (CuSum)

Αλγόριθμοι ανίχνευσης σημείου αλλαγής κατανομής χρησιμοποιήθηκαν αρχικά **στον στατιστικό έλεγχο ποιότητας**. Κατά τον Ghosh (1991) ο πρώτος ακολουθιακός αλγόριθμος δημιουργήθηκε από τον Shewhart (1931), γνωστός και ως **διάγραμμα ελέγχου του Shewhart**. Σε γενικές γραμμές τα διαγράμματα ελέγχου χρησιμοποιούνται για την επιτήρηση διεργασιών, κυρίως ως προς ένα συνεχές χαρακτηριστικό. Συνήθως εξετάζεται **η μέση**

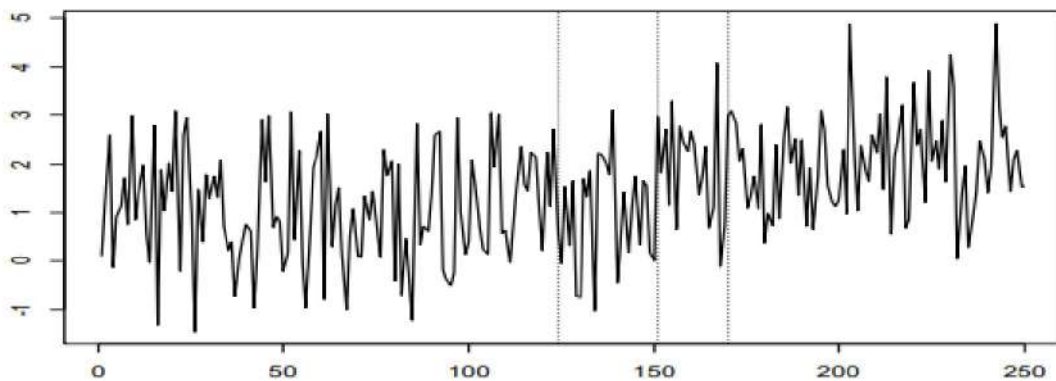
τιμή της κατανομής του χαρακτηριστικού. Σε κάθε διάγραμμα ελέγχου απεικονίζεται μία στατιστική συνάρτηση και τις περισσότερες φορές, **δύο παράλληλες γραμμές**, υποδεικνύοντας ότι εάν η τιμή της συνάρτησης βρεθεί πέρα από το διάστημα που ορίζεται από αυτές τις γραμμές, θα υπάρξει **ένδειξη μετατόπισης του μέσου**.



Εικόνα 3.4: Παράδειγμα Διαγράμματος Shewhart(Πηγή: Analyze-it)

Πολλοί άλλοι αλγόριθμοι κατασκευάστηκαν με σκοπό την ανίχνευση μετατόπισης μέσου ή διασποράς ή και των δύο, οι οποίοι επεκτάθηκαν και σε εφαρμογές πέρα από τον έλεγχο ποιότητας, όπως για παράδειγμα στην στατιστική επεξεργασία σήματος, στην επιτήρηση ιατρικών διαδικασιών, στα συστήματα πλοήγησης κ.α. Εδώ αξίζει να επισημανθούν κάποιες σημαντικές διαφορές. Υπάρχουν δύο είδη στατιστικής ανίχνευσης. Κατά το πρώτο έχουν συλλεχθεί όλα τα δεδομένα και μετά αναζητείται τότε και αν συνέβη κάποια αλλαγή στην κατανομή του χαρακτηριστικού που εξετάζεται. Η διαδικασία αυτή ονομάζεται **offline**. Στο δεύτερο είδος, οι παρατηρήσεις συλλέγονται μία-μία και η ανίχνευση γίνεται εκείνη την χρονική στιγμή. Η διαδικασία αυτή ονομάζεται **online** και αποτελεί μία από τις δημοφιλέστερες εφαρμογές της ακολουθιακής στατιστικής. Ενεργό ερευνητικό πεδίο είναι η **αντιμετώπιση προβλημάτων ταχύτατης ανίχνευσης (quickest detection)**, όπου εξετάζεται η ύπαρξη και η μορφή των **στατιστικών συναρτήσεων** και των **κανόνων απόφασης** που θα ικανοποιούν ορισμένα **κριτήρια βελτιστοποίησης**, για τον κατά το δυνατόν ταχύτερο εντοπισμό μικρής μετατόπισης του μέσου.

Για να περιγραφεί όλη η διαδικασία, θα χρησιμοποιούνται κάποιες βασικές έννοιες από το στατιστικό έλεγχο ποιότητας. Έστω αρχικά ότι εξετάζεται εάν υπάρχει μετατόπιση του μέσου της κατανομής. Θέτοντας ένα κατώφλι (ή δύο ανάλογα με το αν επιζητείται να ελεγχθεί και μείωση του μέσου) και όταν η τιμή της στατιστικής συνάρτησης περάσει αυτήν την τιμή, τότε υπάρχει ένδειξη «συναγερμού» (alarm) ότι ο μέσος της κατανομής άλλαξε, σύμφωνα με την κλασσική ορολογία (π.χ. βλ. Αντζουλάκος, 2012). Διακρίνεται η περίπτωση να υπάρξει συναγερμός ενώ δεν έχει αλλάξει η μέση τιμή της κατανομής, δηλαδή να υπάρξει ένδειξη εσφαλμένου συναγερμού (false alarm ή σφάλμα τύπου I), και να μην υπάρξει συναγερμός όταν στην πραγματικότητα έχει υπάρξει μετατόπιση του μέσου (σφάλμα τύπου II). Στην δεύτερη περίπτωση, ενδεχομένως ο αλγόριθμος να εντοπίσει την αλλαγή αλλά σε προχωρημένο στάδιο. Άρα τότε θα υπάρχει κάποια καθυστέρηση στον εντοπισμό του σημείου αλλαγής. Για παράδειγμα παρατίθεται το παρακάτω σχήμα.



Εικόνα 3.5: Μετατόπιση μέσου κατανομής

Τα πρώτα 150 σημεία προέρχονται από την κατανομή $N(1,1)$ ενώ τα υπόλοιπα 100 από την κατανομή $N(2,1)$. Εάν οι παράμετροι του αλγορίθμου και κυρίως το κατώφλι, ήταν τέτοια ώστε ο αλγόριθμος να ανιχνεύει όσο το δυνατόν γρηγορότερα κάποια αλλαγή, τότε θα ήταν εφικτό να υπήρχε **ένδειξη συναγερμού για αύξηση της μέσης τιμής**, στο σημείο από το που διέρχεται η πρώτη κάθετος αντί για την δεύτερη (δηλαδή αφού δεν έχει αλλάξει η κατανομή, πρόκειται ουσιαστικά για εσφαλμένο συναγερμό). Εάν πάλι, ο

αλγόριθμος ήταν τροποποιημένος ώστε να έχει μικρό αριθμό εσφαλμένων συναγερμών, τότε θα μπορούσε να μην εντόπιζε την αλλαγή κατανομής ικανοποιητικά γρήγορα, με αποτέλεσμα να υπήρχε καθυστέρηση αρκετών σημείων, δηλαδή να εντοπιζόταν η αλλαγή κατανομής στο σημείο που διέρχεται η τρίτη κάθετος.

Σκοπός των μεθόδων της ακολουθιακής ανάλυσης είναι να ελαχιστοποιηθεί η μέση καθυστέρηση ανίχνευσης (mean delay) όταν ο αναμενόμενος χρόνος μεταξύ εσφαλμένων συναγερμών (mean time between false alarms) είναι σταθερός.

Αλγόριθμος CUSUM

Έστω i.i.d. τ.μ. $(X_n)_{n \geq 1}$ όπου $X_i \sim f(x|\theta)$ με μέση τιμή θ και διασπορά $0 < \sigma < \infty$. Εξετάζεται αρχικά ο έλεγχος της υπόθεσης:

$$H_0: \theta = \theta_0 \text{ έναντι της } H_1: \theta = \theta_1 \text{ (} \theta_1 > \theta_0 \text{)} \quad (13)$$

όπου $\theta_0, \theta_1 \in \Theta \subseteq \mathbb{R}$. Το μέγεθος του δείγματος σύμφωνα με την κλασική θεωρία των Neyman-Pearson (1933) είναι σταθερό και γνωστό εκ των προτέρων.

Ο αλγόριθμος σωρευτικών αθροισμάτων (cumulative sum-CUSUM) προτάθηκε από τον Page (1954). Ο CUSUM σε αντίθεση με τα κλασσικά διαγράμματα Shewhart λαμβάνει υπ' όψιν του τις προηγούμενες παρατηρήσεις μέσω ενός αθροίσματος.

$$\{x_1, x_1 + x_2, x_1 + x_2 + x_3, \dots, \sum_{i=1}^n x_i\} \quad (14)$$

Λαμβάνοντας δείγμα μεγέθους ένα κάθε φορά, από την κανονική κατανομή $N(\theta, \sigma^2)$ και κάνοντας τον έλεγχο υπόθεσης:

Έστω αρχικά ότι τα θ_0 και θ_1 καθώς και η διακύμανση σ^2 είναι γνωστά, και έστω τα αθροίσματα:

$$S_n = \sum_{i=1}^n (x_i - \theta_0). \quad (15)$$

Όσο ισχύει η H_0 αναμένεται τα σημεία (n, S_n) να κινούνται γύρω από τον άξονα x γιατί $E(S_n) = 0$. Εάν στο σημείο με τετμημένη τ_0 η μέση τιμή γίνει $\theta_1 > \theta_0$ τότε:

$$E(S_n) = E\left(\sum_{i=1}^{\tau_0-1} (X_i - \theta_0) + \sum_{i=\tau_0}^n (X_i - \theta_0)\right) = (\theta_1 - \theta_0)(n - \tau_0) \quad (16)$$

Άρα τα σημεία (n, S_n) για $n > \tau_0$ θα βρίσκονται γύρω από ευθεία με κλίση $\lambda = (\theta_1 - \theta_0)$. Σημειώνεται ότι το τ_0 ονομάζεται σημείο αλλαγής κατανομής (change point).

3.2.3 Μεταβλητότητα Συμπεριφοράς

Ο δείκτης **μεταβλητότητας συμπεριφοράς (Volatility)** αποτελεί ένα μαθηματικό-στατιστικό εργαλείο το οποίο χρησιμοποιείται κατά κύριο λόγο για την παρατήρηση του τρόπου με τον οποίο οι οδηγοί **μεταβαίνουν από μια κατάσταση οδηγικής συμπεριφοράς σε μια άλλη** (π.χ. από ασφαλή κατάσταση οδήγησης σε επιθετική ή επικίνδυνη ή και το αντίθετο) από ταξίδι σε ταξίδι. Πιο συγκεκριμένα, διενεργείται ανάλυση για κάθε ταξίδι κάθε οδηγού ξεχωριστά προκειμένου να καθοριστεί η μέση οδηγική συμπεριφορά κάθε οδηγού, εξετάζοντας ταυτόχρονα αν αυτή χειροτερεύει ή καλυτερεύει από ταξίδι σε ταξίδι και αν γενικά η συνολική συμπεριφορά του παρουσιάζει σταθερότητα ή μεταβλητότητα.

Για να υπολογιστεί η μεταβλητότητα, πρέπει πρώτα να υπολογιστεί ο **λόγος Κέρδος / Απώλεια** του κάθε οδηγού. Ο λόγος αυτός ορίζεται ότι αντιστοιχεί στο κέρδος (=βελτίωση) ή στην απώλεια αντίστοιχα, που ένας οδηγός λαμβάνει **σύμφωνα με τον τρόπο που μετακινείται μεταξύ διαδοχικών ταξιδιών**. Έστω k το χαρακτηριστικό που εξετάζεται (π.χ. harsh acceleration events, harsh braking events, time mobile usage, time speeding κλπ.), i ο αριθμός του χρήστη ($i = 1, 2, 3, \dots, N$) και t ο αριθμός των ταξιδιών $t = \{1, 2, 3, \dots, n\}$ κάθε οδηγού. Ο λόγος Κέρδος / Απώλεια για κάθε ταξίδι δίνεται από τη σχέση:

$$r_{t,i} = \ln \left(\frac{k_{t,i}}{k_{t-1,i}} \right) \quad (17)$$

Αυτός ο δείκτης έχει **θετική αξία**, όταν ο οδηγός **«βελτιώνει» τη συμπεριφορά του** (π.χ. μείωση του αριθμού των απότομων επιβραδύνσεων ανά χιλιόμετρο στο τελευταίο ταξίδι σε σύγκριση με το προηγούμενο) έναντι του προηγούμενου ταξιδιού, ενώ οι **αρνητικές τιμές** δείχνουν ότι ο οδηγός **έχει κινηθεί προς μια χειρότερη συμπεριφορά** (π.χ. αύξηση του αριθμού των απότομων επιβραδύνσεων ανά χιλιόμετρο στο τελευταίο ταξίδι σε σύγκριση με το προηγούμενο).

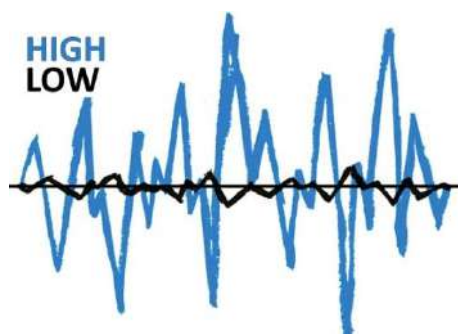
Στη συνέχεια, το μέτρο μεταβλητότητας της συμπεριφοράς του οδηγού υπολογίζεται ως **η τυπική απόκλιση του λόγου Κέρδους / Απώλεια**, προκειμένου να ανιχνευθεί **σταθερότητα** του τρόπου μετακίνησης του οδηγού μεταξύ των διαφορετικών ταξιδιών:

$$\text{Volatility} = \sqrt{\frac{\sum_{t=1}^n (r_{t,i} - \bar{r}_i)^2}{n-1}} \quad (18)$$

όπου $r_{t,i}$ είναι ο λόγος Κέρδος / Απώλεια για κάθε ταξίδι t κάθε χρήστη i , $\bar{r}_i$ μέση τιμή του λόγου Κέρδος / Απώλεια για τον οδηγό i και n ο αριθμός των ταξιδιών.

Για να υπολογιστεί την μέση τιμή $\bar{r}$ του λόγου Κέρδος / Απώλεια, θα πρέπει να καθοριστεί ο αριθμός των παρατηρήσεων που θα χρησιμοποιηθεί για τον σκοπό αυτό. Για παράδειγμα, αν χρησιμοποιηθούν όλα τα ταξίδια ενός οδηγού για τον υπολογισμό του $\bar{r}$ τότε θα υπολογίζεται το volatility συγκρινόμενο με όλο το δείγμα. Όμως διαφορετικοί οδηγοί έχουν διαφορετικό πλήθος ταξιδιών (γενικά, αλλά και σε κάθε κατηγορία χρονικής διάρκειας) και συνεπώς δεν θα ήταν δόκιμο κάτι τέτοιο. Αντιθέτως, επιλέχθηκε να χρησιμοποιηθεί ένα «**κινούμενο παράθυρο**» το οποίο ήταν σταθερό και ίσο με **20 ταξίδια**. Ο αριθμός αυτός ορίστηκε κάνοντας την εύλογη **παραδοχή** ότι κανένας οδηγός και κανένα χαρακτηριστικό δεν συγκλίνει νωρίτερα από τα 20 ταξίδια παρατήρησης, και άρα ένας οδηγός χρειάζεται να παρατηρηθεί σίγουρα για περισσότερα από 20 ταξίδια. Επομένως, η πρώτη τιμή του volatility του χρήστη i αφορά στη σύγκριση του πρώτου ταξιδιού ($r_{1,i}$) με την μέση τιμή ($\bar{r}_i$) των 20 επόμενων (συμπεριλαμβανομένου και αυτού) κ.ο.κ.

Οι τιμές του volatility είναι πάντα θετικές. Όσο μεγαλύτερη είναι η τιμή του volatility, τόσο πιο μεταβαλλόμενη είναι η συμπεριφορά του οδηγού, και το αντίθετο, όσο μικρότερη είναι η τιμή του, τόσο πιο σταθερή είναι η οδηγική συμπεριφορά. Για να γίνει ευκολότερα κατανοητό αυτό, παρατίθεται η παρακάτω εικόνα.



Εικόνα 3.6: Παράδειγμα υψηλής και χαμηλής μεταβλητότητας

Με την μπλε γραμμή παριστάνεται ένα μέγεθος που παρουσιάζει μεγάλες και απότομες διακυμάνσεις, έχει δηλαδή μεγάλη μεταβλητότητα (μεγάλο volatility).

Σε αντίθεση με αυτό, η μαύρη γραμμή παριστάνει ένα μέγεθος πιο σταθερό, με μικρή μεταβλητότητα(μικρό volatility).

Συνεπώς, όταν οι τιμές του volatility φτάσουν να είναι αρκετά χαμηλές, ενώ παράλληλα διαδοχικές τιμές του volatility δεν αποκλίνουν σημαντικά η μία από την άλλη, τότε μπορεί να ειπωθεί ότι αυτή είναι μια ένδειξη σταθεροποίησης της συμπεριφοράς ενός οδηγού.

3.2.4 Κινητοί Μέσοι Όροι

Ο κινητός μέσος όρος είναι ένας **μέσος όρος ενός κινούμενου συνόλου τιμών**, ο οποίος υπολογίζεται για έναν **συγκεκριμένο αριθμό ημερών**. Ο κινητός μέσος όρος **διευκολύνει την οπτικοποίηση των τάσεων**, καθώς εξαλείφει - ή τουλάχιστον ελαχιστοποιεί - τον καθημερινό στατιστικό θόρυβο. Ο κινητός μέσος όρος αποτελεί ένα από τα παλαιότερα εργαλεία της Τεχνικής Ανάλυσης και είναι μια δημοφιλής μέθοδος αναγνώρισης και αξιοποίησης των τάσεων.

Οι κινητοί μέσοι όροι χρησιμοποιούνται κατά κόρον στους χρηματιστηριακούς δείκτες και αποτελούνται από τον μέσο όρο των κλεισιμάτων μιας μετοχής με συγκεκριμένο χρονικό ορίζοντα. Ο κυριότερος λόγος που χρησιμοποιούνται στις μαθηματικές μεθόδους συναλλαγών είναι ότι προσφέρουν μια αντικειμενική άποψη της κίνησης των τιμών. Παρέχουν στον αναλυτή μια εικόνα της κατεύθυνσης των τιμών και της πορείας τους στο χρόνο, μειώνοντας το φόβο και αποφεύγοντας την υπερβολική δραστηριότητα. Επίσης, κάποιες φορές αυτές οι μέθοδοι προσφέρουν προειδοποιητικές ενδείξεις ως προς τις κινήσεις των τιμών, πριν την υλοποίησή τους στα πλαίσια της αγοράς. Οι κινητοί μέσοι όροι χρησιμοποιούνται επομένως για:

- να περιορίσουν τις μικρές και μη σημαντικές διακυμάνσεις στην τιμή.
- να βοηθούν στον προσδιορισμό της επικρατούσας τάσης.

Υπάρχουν πολλά ήδη κινητών μέσων όρων. Οι πιο συχνά χρησιμοποιούμενοι είναι οι εξής:

- ο Απλός Κινητός Μέσος Όρος (Simple Moving Average)
- ο Σταθμισμένος Κινητός Μέσος Όρος (Linearly Weighted Moving Average)
- ο Εκθετικός Κινητός Μέσος Όρος (Exponentially Smoothed Moving Average)

Απλός Κινητός Μέσος Όρος

Ο απλός κινητός μέσος όρος (**Simple Moving Average – SMA**) είναι ένας αριθμητικός μέσος όρος δεδομένων τιμών. Είναι ο πιο απλός και ευρέως διαδεδομένος κινητός μέσος όρος. Δεν πρέπει να συγχάεται με τον γενικό μέσο όρο, καθώς ο κινητός μέσος όρος αφορά ένα υποσύνολο μιας σειράς παρατηρήσεων, το οποίο κάθε φορά αγνοεί μια παλαιότερη παρατήρηση και συμπεριλαμβάνει την αμέσως νεότερη. Δεδομένης μιας σειράς παρατηρήσεων και ενός σταθερού μεγέθους υποσυνόλου, το πρώτο στοιχείο του κινητού μέσου λαμβάνεται λαμβάνοντας τον μέσο όρο του αρχικού σταθερού υποσυνόλου των σειρών αριθμών. Στη συνέχεια, **το υποσύνολο τροποποιείται με "μετατόπιση προς τα εμπρός"** δηλαδή, εξαιρώντας τον πρώτο αριθμό της σειράς και συμπεριλαμβάνοντας την επόμενη τιμή στο υποσύνολο. Υπολογίζεται προσθέτοντας την τιμή κάθε διαστήματος και διαιρώντας το σύνολο με τον αριθμό των διαστημάτων που καλύπτει ο κινητός μέσος όρος.

$$SMA_M = \frac{1}{n} \sum_{M=1}^n p_M \quad (19)$$

Για παράδειγμα, για να υπολογιστεί κινητός μέσος όρος 25 ημερών των τιμών κλεισίματος ενός τίτλου στο χρηματιστήριο, προστίθενται τιμές κλεισίματος των τελευταίων 25 ημερών του τίτλου και στη συνέχεια

διαίρουνται διά 25. Θα πρέπει να ληφθεί υπόψιν ότι η τιμή που προστίθεται και αυτή που αφαιρείται κάθε φορά, μπορούν να επηρεάσουν σημαντικά τον μέσο όρο και ότι όλα τα δεδομένα έχουν την ίδια βαρύτητα.

Κύριο πλεονέκτημα του SMA είναι ότι πρόκειται για έναν από τους καλύτερους τρόπους εκτίμησης της δύναμης μιας μακροπρόθεσμης τάσης και της πιθανότητας αντιστροφής. Όταν ένας κινητός μέσος όρος είναι ανοδικός και η τιμή τον υπερβαίνει, ο τίτλος έχει ανοδική τάση. Αντιστρόφως, ένας καθοδικός μέσος όρος με την τιμή να βρίσκεται σε χαμηλότερο επίπεδο υποδεικνύει καθοδική τάση. Στα μειονεκτήματά του ανήκει ότι προσφέρει ενδείξεις **μετά την έναρξη** μιας νέας κίνησης, ενός συμβάντος ή μιας τάσης **καιόχι πριν**.

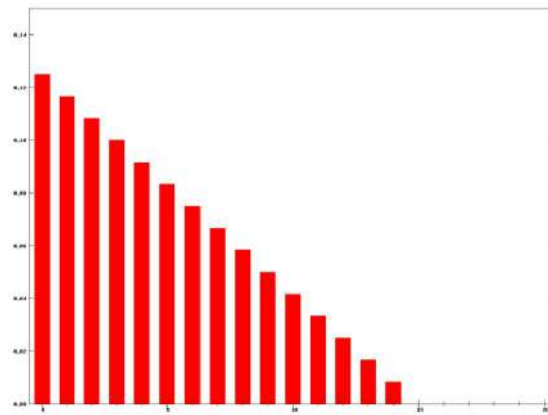
Σταθμισμένος Κινητός Μέσος Όρος (WMA)

Ο **σταθμισμένος κινητός μέσος όρος (WMA)** είναι ένας μέσος όρος που έχει **πολλαπλασιαστικούς παράγοντες** που δίνουν **διαφορετικά βάρη** στα δεδομένα **σε διαφορετικές θέσεις** στο κινούμενο παράθυρο του δείγματος.

Μαθηματικά, ο κινητός μέσος όρος είναι η συνέλιξη των σημείων αναφοράς με σταθερή συνάρτηση σταθμίσεως. Για τον υπολογισμό του προστίθενται οι τιμές κλεισίματος **πολλαπλασιασμένες** με την **στάθμισή τους** και το αποτέλεσμα διαιρείται με το άθροισμα των σταθμίσεων. Η στάθμιση της πιο πρόσφατης τιμής εξαρτάται από την εκάστοτε χρονική περίοδο του κινητού μέσου όρου. Όσο πιο μικρή είναι η χρονική περίοδος του EMA, τόσο μεγαλύτερο βάρος δίνεται στην πιο πρόσφατη τιμή. Στην τεχνική ανάλυση των οικονομικών στοιχείων, ο σταθμισμένος κινητός μέσος όρος (WMA) έχει την ειδική σημασία των βαρών που μειώνονται στην αριθμητική εξέλιξη. Για τον υπολογισμό του WMA για n -ημέρες η τελευταία ημέρα (πιο πρόσφατη) έχει βάρος n , η προ-τελευταία $n-1$, κλπ. μέχρι την παλαιότερη με βάρος 1.

$$WMA_M = \frac{n \cdot p_M + (n-1) \cdot p_{M-1} + \dots + 2 \cdot p_{M-n+2} + 1 \cdot p_{M-n+1}}{n + (n-1) + \dots + 2 + 1} \quad (20)$$

Το παρακάτω γράφημα δείχνει τον τρόπο μείωσης του βάρους, από το υψηλότερο βάρος για τα πιο πρόσφατα σημεία αναφοράς, μέχρι το μηδέν. Μπορεί να συγκριθεί με τα βάρη στον εκθετικό κινούμενο μέσο όρο που ακολουθεί.



Εικόνα 3.7: Παράδειγμα μείωσης των συντελεστών βάρους για WMA(Πηγή: Wikipedia)

Εκθετικός Κινητός Μέσος Όρος (EMA)

Ο **εκθετικός κινητός μέσος όρος (EMA)**, γνωστός και ως εκθετικά σταθμισμένος κινητός μέσος όρος (EWMA), είναι ένας **σταθμισμένος μέσος όρος** των δεδομένων μιας τιμής, ο οποίος εφαρμόζει συντελεστές στάθμισης οι οποίοι μειώνονται εκθετικά, δηλαδή **δίνει μεγαλύτερο βάρος στα πιο πρόσφατα δεδομένα**. Η στάθμιση για κάθε παλαιότερη βάση μειώνεται εκθετικά, χωρίς να φθάσει ποτέ στο μηδέν. Για τον υπολογισμό του προστίθενται οι τιμές κλεισίματος πολλαπλασιασμένες με την στάθμισή τους και το αποτέλεσμα διαιρείται με το άθροισμα των σταθμίσεων. Η στάθμιση της πιο πρόσφατης τιμής εξαρτάται από την εκάστοτε χρονική περίοδο του κινητού μέσου όρου. Όσο πιο μικρή είναι η χρονική περίοδος του EMA, τόσο μεγαλύτερο βάρος δίνεται στην πιο πρόσφατη τιμή.

Ο EMA ορίζεται με δύο τρόπους: ως **EMA βασιζόμενος σε ποσοστά**, όπου ο αναλυτής ορίζει το ποσοστιαίο βάρος της τιμής της τελευταίας περιόδου, ή ως **EMA βασιζόμενος σε χρονική περίοδο**, όπου ο αναλυτής ορίζει τη

διάρκεια του EMA και το βάρος κάθε χρονικής περιόδου υπολογίζεται μέσω συγκεκριμένου τύπου. Περισσότερο χρησιμοποιείται ο δεύτερος τύπος.

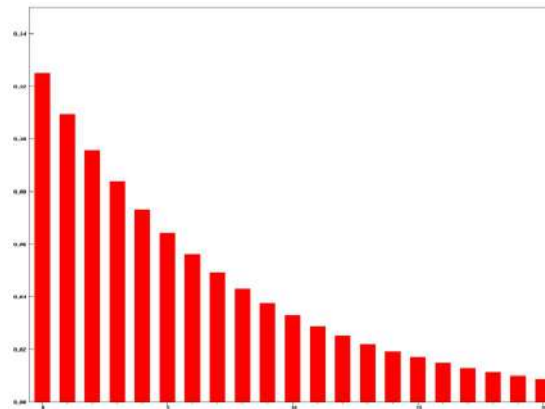
Σε αντίθεση με τον Απλό Κινητό Μέσο Όρο και τον Σταθμισμένο Κινητό Μέσο Όρο, στον υπολογισμό του EMA **χρησιμοποιούνται όλες οι προηγούμενες τιμές** των εκάστοτε δεδομένων. Παρόλο που ο αντίκτυπος των παλιότερων δεδομένων μικραίνει με το πέρασμα του χρόνου, δεν εξαλείφεται ποτέ. Το τελευταίο έχει ως αποτέλεσμα ακόμη και αν η τιμή μείνει αμετάβλητη για n bars, ο SMA & EMA να διαφέρουν στην τιμή. Αυτό ισχύει ανεξάρτητα από την καθορισμένη χρονική περίοδο του EMA. Οι επιπτώσεις των παλιότερων δεδομένων εξασθενούν σύντομα για τους EMA μικρότερης διάρκειας σε σύγκριση με τους EMA μεγαλύτερης διάρκειας, αλλά δεν παύουν ποτέ να υφίστανται. Αναφορικά, ο EMA υπολογίζεται ως εξής:

$$EMA_{today} = \frac{p_1 + (1-a)p_2 + (1-a)^2 p_3 + (1-a)^3 p_4 + \dots}{1 + (1-a) + (1-a)^2 + ((1-a)^3 + \dots)} \quad (21)$$

όπου p_1 είναι η τιμή για σήμερα, p_2 είναι η τιμή για χθες κ.ο.κ. Ο συντελεστής a αντιπροσωπεύει τον βαθμό της μείωσης του βάρους, έναν σταθερό παράγοντα εξομάλυνσης μεταξύ 0 και 1. Μια υψηλότερη τιμή του a μειώνει τον αντίκτυπο των παλαιότερων παρατηρήσεων γρηγορότερα

Κύριο Πλεονέκτημα του EMA συγκριτικά με τον SMA είναι ότι, επειδή ο EMA δίνει μεγαλύτερο βάρος στις πιο πρόσφατες παρατηρήσεις, επιτρέπει στους τεχνικούς επενδυτές να αντιδρούν γρηγορότερα στις πρόσφατες αλλαγές των τιμών.

Το παρακάτω γράφημα δείχνει ένα παράδειγμα του τρόπου μείωσης του βάρους.

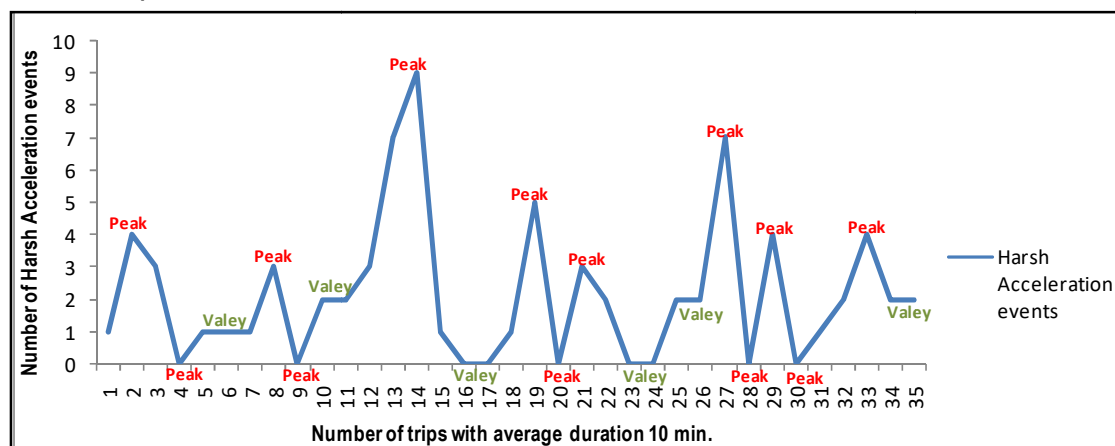


Εικόνα 3.8: Παράδειγμα μείωσης των συντελεστών βάρους για EMA(Πηγή: Wikipedia)

3.2.5 Τοπικά Ακρότατα

Η ανάλυση δεδομένων χρονοσειρών, χαρακτηρίζεται σε έντονο βαθμό από το μεγάλο πλήθος διακυμάνσεων στις τιμές της εκάστοτε εξεταζόμενης μεταβλητής. Αυτές οι διακυμάνσεις ταξινομούνται κυρίως σε κορυφές (peaks) και κοιλάδες (valleys) όπως φαίνεται στο παρακάτω σχήμα.

Μια κορυφή ή μια κοιλάδα είναι ένα σημαντικό γεγονός μέσα σε μια μαθηματική συνάρτηση ή μια χρονοσειρά, καθώς αποτελεί ένα σημείο όπου το γράφημα λειτουργίας αλλάζει από την αυξανόμενη (μειούμενη) συμπεριφορά στη μειούμενη (αυξανόμενη) συμπεριφορά. Ο προσδιορισμός αυτής της συμπεριφοράς είναι σημαντικός προκειμένου να διεξαχθεί ανάλυση των δεδομένων.



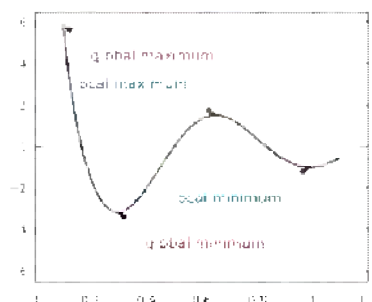
Εικόνα 3.9: Απόσπασμα χρονοσειράς γεγονότων απότομων επιταχύνσεων ενός τυχαίου χρήστη για ταξίδια μέσης διάρκειας 10 λεπτών

Μαθηματικοί Ορισμοί και Διατυπώσεις

Στην μαθηματική ανάλυση, το μέγιστο και το ελάχιστο μιας συνάρτησης (ακρότατα) είναι οι μεγαλύτερες και μικρότερες τιμές της συνάρτησης είτε μέσα σε ένα δεδομένο εύρος (το τοπικό ή σχετικό ακρότατο) ή σε ολόκληρο το πεδίο μιας συνάρτησης (το ολικό ακρότατο).

- ❖ Μια συνάρτηση πραγματικής τιμής f ορισμένη σε ένα πεδίο X παρουσιάζει **ολικό μέγιστο** στο σημείο x^* αν $f(x^*) \geq f(x)$, $\forall x \in X$.
- ❖ Αντίστοιχα, μια συνάρτηση πραγματικής τιμής f ορισμένη σε ένα πεδίο X παρουσιάζει **ολικό ελάχιστο** στο σημείο x^* αν $f(x^*) \leq f(x)$, $\forall x \in X$.
- ❖ Εάν το πεδίο X είναι ένας μετρικός χώρος τότε το f λέγεται ότι έχει ένα **τοπικό (ή σχετικό) μέγιστο** σημείο στο σημείο x^* εάν υπάρχει κάποιο $\varepsilon > 0$ τέτοιο ώστε $f(x^*) \geq f(x)$ για όλα τα $x \in X$ στην απόσταση ε από το x^* .
- ❖ Παρομοίως, η συνάρτηση έχει ένα **τοπικό (ή σχετικό) ελάχιστο** σημείο στο x^* εάν $f(x^*) \leq f(x)$ για όλα τα $x \in X$ στην απόσταση ε από το x^* .

Ένας παρόμοιος ορισμός μπορεί να χρησιμοποιηθεί όταν το X είναι ένας τοπολογικός χώρος, αφού ο ορισμός που μόλις δόθηκε μπορεί να αναδιατυπωθεί σε όρους περιοχών γύρω από το x^* .



Εικόνα 3.10: Παράδειγμα τοπικών και ολικών ακροτάτων (Πηγή: Wikipedia)

Η εύρεση ολικών μεγίστων και ελάχιστων είναι ο στόχος της μαθηματικής βελτιστοποίησης. Εάν μια συνάρτηση είναι συνεχής σε ένα κλειστό διάστημα, τότε από το θεώρημα της ακραίας τιμής υπάρχουν μέγιστα και ελάχιστα σημεία. Επιπλέον, ένα μέγιστο (ή ελάχιστο) πρέπει είτε να είναι τοπικό μέγιστο (ή ελάχιστο) στο εσωτερικό του τομέα, είτε πρέπει να βρίσκεται στο όριο του τομέα. Έτσι μια μέθοδος για την εξεύρεση ενός ολικού μέγιστου (ή ελάχιστου) είναι να εξεταστούν όλα τα τοπικά μέγιστα (ή τα ελάχιστα) στο εσωτερικό και επίσης να δούμε τα μέγιστα (ή τα ελάχιστα) των σημείων στο όριο και να επιλεγεί το μεγαλύτερο (ή μικρότερο) από αυτά.

Το τοπικό ακρότατο των παραγωγίσιμων συναρτήσεων μπορεί να βρεθεί από το **θεώρημα Fermat**, το οποίο αναφέρει ότι αυτά πρέπει να εμφανίζονται σε «κρίσιμα σημεία». Μπορεί κανείς να διακρίνει εάν ένα κρίσιμο σημείο είναι τοπικό μέγιστο ή τοπικό ελάχιστο χρησιμοποιώντας τη δοκιμή της πρώτης παραγώγου, της δεύτερης παραγώγου ή τη δοκιμή παράγωγου υψηλότερης τάξης, με δεδομένη επαρκή παραγωγήση.

4. ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΣΤΟΙΧΕΙΩΝ

4.1 Εισαγωγή

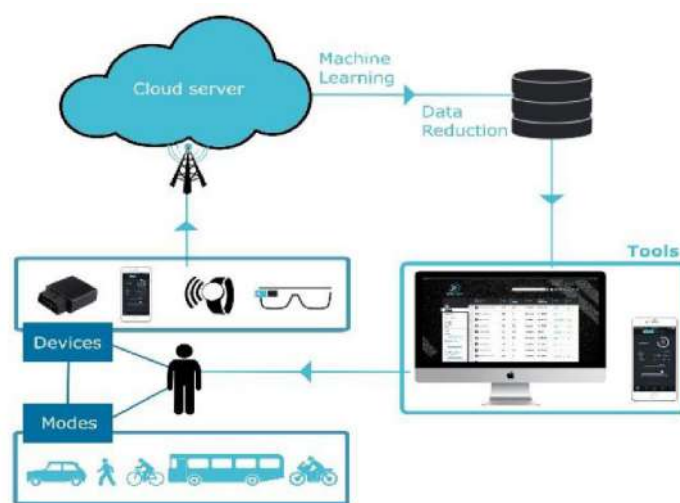
Το κεφάλαιο αποτελείται από δύο υπό-κεφάλαια. Στο εδάφιο που αναφέρεται στη συλλογή στοιχείων, περιγράφεται το πείραμα που πραγματοποιήθηκε για τα στοιχεία που συλλέχθηκαν. Στο δεύτερο εδάφιο το οποίο αφορά στην εφαρμογή της προτεινόμενης μεθοδολογικής προσέγγισης, παρατίθενται συγκεντρωτικά διαγράμματα στα οποία παρουσιάζονται τα χαρακτηριστικά του δείγματος, ακολουθούμενα από τον απαραίτητο σχολιασμό.

4.2 Συλλογή Στοιχείων

Σύμφωνα με τη βιβλιογραφική ανασκόπηση, προκύπτει ότι έχουν πραγματοποιηθεί αρκετές έρευνες σχετικά με την παρακολούθηση και καταγραφή των χαρακτηριστικών της οδηγικής συμπεριφοράς, τόσο σε πραγματικές συνθήκες οδήγησης, όσο και σε προσομοιωτές. Ωστόσο, στις περιπτώσεις **πειραμάτων σε πραγματικές συνθήκες οδήγησης**, η μεθοδολογία συλλογής των στοιχείων περιλάμβανε συνήθως μια συσκευή που βρισκόταν εντός του οχήματος και ήταν συνδεδεμένη με το όχημα. Παραδείγματα τέτοιων συσκευών είναι το **Driving Data Recorder DDR** (Ohta, Tohru, and Shouji Nakajima, 1994), μια συσκευή μικρών διαστάσεων που συνδεόταν με τη μηχανή του αυτοκινήτου ή το **On Board Diagnostics (OBD)** (Zaldivar, Jorge, et al., 2011) το οποίο έχει ως κύριο στόχο την ανίχνευση μηχανικών προβλημάτων στο όχημα, αλλά χρησιμοποιήθηκε για την καταγραφή δεδομένων. Στην παρούσα διπλωματική εργασία η συλλογή των στοιχείων γίνεται μέσω αισθητήρων από κινητό τηλέφωνο-smartphone.

Πιο αναλυτικά, εφαρμόστηκε ένα καινοτόμο σύστημα συλλογής δεδομένων, μέσα από την εξατομικευμένη καταγραφή της οδηγικής συμπεριφοράς σε πραγματικό χρόνο χρησιμοποιώντας αισθητήρες συσκευών smartphone. Για τον σκοπό αυτό, αναπτύχθηκε μια εφαρμογή για έξυπνα κινητά τηλέφωνα

(smartphones) από την εταιρεία **OSeven Telematics**, η οποία χρησιμοποιήθηκε για την καταγραφή της οδηγικής συμπεριφοράς των συμμετεχόντων χρηστών. Στο πείραμα (smartphone naturalistic driving experiment) συμμετείχαν **68 οδηγοί** σε χρονικό διάστημα **7 μηνών** και δημιουργήθηκε μια μεγάλη βάση δεδομένων με χιλιάδες (21.610) διαδρομές. Οι μετρήσεις πραγματοποιήθηκαν τόσο σε αστικό όσο και σε υπεραστικό περιβάλλον και αυτοκινητόδρομο. Τα βήματα που ακολουθούνται κάθε φορά που καταγράφεται μια διαδρομή-ταξίδι απεικονίζονται στην παρακάτω εικόνα 4.1 και περιγράφονται αναλυτικά στην επόμενη παράγραφο.



Εικόνα 4.1 Διάγραμμα διαχείρισης δεδομένων (Πηγή: OSeven)

4.2.1 Μεθοδολογία καταγραφής, συλλογής, επεξεργασίας και οπτικοποίησης δεδομένων

Η OSeven έχει αναπτύξει ένα ολοκληρωμένο σύστημα για την καταγραφή, την συλλογή, την αποθήκευση, την αξιολόγηση και την παρουσίαση των δεδομένων της οδηγικής συμπεριφοράς χρησιμοποιώντας εφαρμογές έξυπνων κινητών τηλεφώνων (smartphone) και προηγμένους αλγόριθμους Machine Learning. Αυτή η καινοτομία που εφαρμόζεται στον τρόπο συλλογής και ανάλυσης δεδομένων σε μεγάλη κλίμακα, παρουσιάζει νέες προκλήσεις συγκεντρώνοντας μεγάλο όγκο δεδομένων για ανάλυση κατά τη διάρκεια

αυτής της έρευνας. Το σύστημα που αναπτύχθηκε ενσωματώνει μια διαδικασία συλλογής και διαβίβασης δεδομένων από τα έξυπνα κινητά τηλέφωνα (Smartphones), των οποίων τα κύρια χαρακτηριστικά περιγράφονται στις επόμενες παραγράφους.

Σύστημα Καταγραφής Δεδομένων

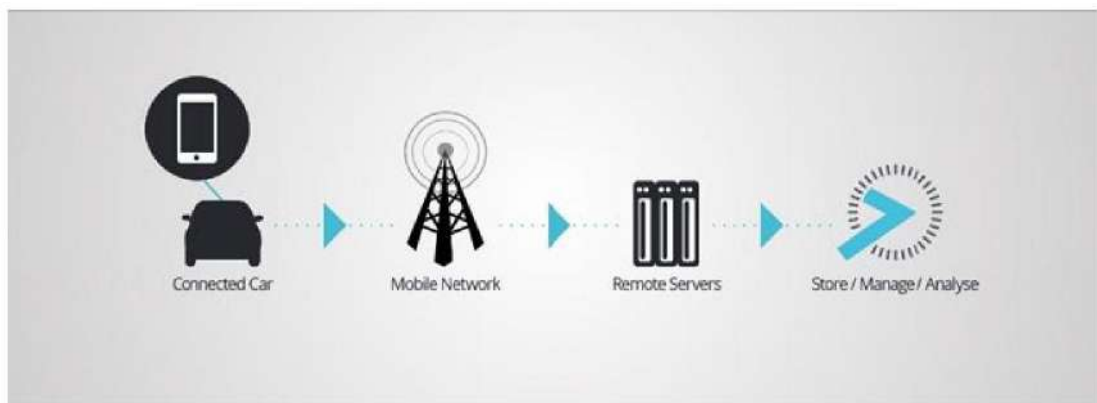
Η καταγραφή των δεδομένων ξεκινάει αυτόματα στις εφαρμογές του κινητού τηλεφώνου όταν αναγνωρίζεται μια κατάσταση οδήγησης και σταματάει αυτόματα όταν αναγνωρίζεται κατάσταση μη οδήγησης. Η εγγραφή κατά τη διάρκεια της οδήγησης μένει σε αναμονή για πέντε λεπτά μετά την διακοπή της κατάστασης της οδήγησης, για να ληφθεί υπόψη το γεγονός ότι ο οδηγός συνεχίζει το ταξίδι του με στάση λίγων λεπτών. Όλες οι επιπλέον πληροφορίες που συλλέγονται μετά το πέρας της διαδρομής απορρίπτονται χρησιμοποιώντας τις τεχνικές μηχανικής μάθησης που περιγράφονται παρακάτω.

Τα καταγεγραμμένα δεδομένα προέρχονται από τους διάφορους ενσωματωμένους αισθητήρες των έξυπνων κινητών τηλεφώνων (smartphones) και αλγόριθμους συγχώνευσης (fusion algorithms) δεδομένων που παρέχονται από το Android (Google) και το iOS (Apple). Η εφαρμογή κινητού τηλεφώνου (mobile App) που έχει αναπτυχθεί καταγράφει τη συμπεριφορά του χρήστη χρησιμοποιώντας τους αισθητήρες της συσκευής και χρησιμοποιείται μια ποικιλία APIs για να διαβαστούν τα δεδομένα των αισθητήρων που έχουν καταγραφεί και να αποθηκευτούν προσωρινά στη βάση δεδομένων του έξυπνου κινητού τηλεφώνου (Smartphone) πριν μεταφερθούν στην κεντρική βάση δεδομένων. Μετά τη μετάδοση των δεδομένων στην κεντρική βάση δεδομένων του συστήματος, όλες οι πληροφορίες διαγράφονται από το κινητό τηλέφωνο. Τα πρωτογενή δεδομένα είτε καταγράφονται από τους αισθητήρες των έξυπνων κινητών τηλεφώνων (Smartphone) είτε παρέχονται από iOS και Android με τη μορφή δεδομένων

συγχώνευσης. Ενδεικτικά, το κινητό τηλέφωνο ενσωματώνει αισθητήρες τεχνολογίας:

- Επιταχυνσιόμετρου (Accelerometer)
- Γυροσκόπιου (Gyroscope)
- Μαγνητόμετρου (Magnetometer)
- GPS (ταχύτητα, πορεία, γεωγραφικό μήκος, γεωγραφικό πλάτος)
- Τα δεδομένα που παρέχονται από iOS και Android:
 - Yaw, Pitch, Roll
 - Γραμμική επιτάχυνση
 - Βαρύτητα

Η συχνότητα της καταγραφής των δεδομένων ποικίλλει ανάλογα με τον τύπο του αισθητήρα, με μέγιστη τιμή 1Hz. Αξίζει να τονιστεί ότι με αυτόν τον τρόπο συλλέγεται ένα μεγάλο πλήθος δεδομένων των χαρακτηριστικών οδήγησης και αυτό πραγματοποιείται μόνο με την χρήση των εφαρμογών του κινητού τηλεφώνου. Το βασικό πλαίσιο λειτουργίας της ροής δεδομένων φαίνεται στο σχήμα που ακολουθεί.



Εικόνα4.2: Σύστημα ροής δεδομένων (Πηγή: OSeven)

Μετάδοση δεδομένων

Μετά την ολοκλήρωση της διαδρομής, η εφαρμογή μεταδίδει όλα τα δεδομένα που έχουν συλλεχθεί, στην **κεντρική βάση δεδομένων** του **OSeven backend office** μέσω ενός κατάλληλου δίαυλου επικοινωνίας, όπως είναι ένα

δίκτυο Wi-Fi ή ένα κυψελοειδές δίκτυο (κατόπιν επιλογής του χρήστη) όπως είναι το δίκτυο 3G / 4G με βάση τις ρυθμίσεις του χρήστη.

Για να επιτευχθεί η διάδραση μεταξύ αυτών των πλευρών, δημιουργήθηκε API το οποίο χρησιμοποιείται για την εισαγωγή των δεδομένων από μια ηλεκτρονική υπηρεσία σε μια εφαρμογή πελάτη. Το API χρησιμοποιείται για τη μεταφορά και τη λήψη δεδομένων μεταξύ συστημάτων, υποστηρίζοντας τη μεταξύ τους διαλειτουργικότητα. Κάνοντας τα δεδομένα προσβάσιμα μέσω του Παγκόσμιου Ιστού με τη χρήση API, καθίσταται δυνατή η υποβολή δεδομένων στη βάση δεδομένων από συσκευές τρίτων καθώς επίσης και την πληροφορία άμεσα διαθέσιμη. Ο συνολικός όγκος δεδομένων για έναν μέσο οδηγό υπολογίζεται σε 50Mb / μήνα.

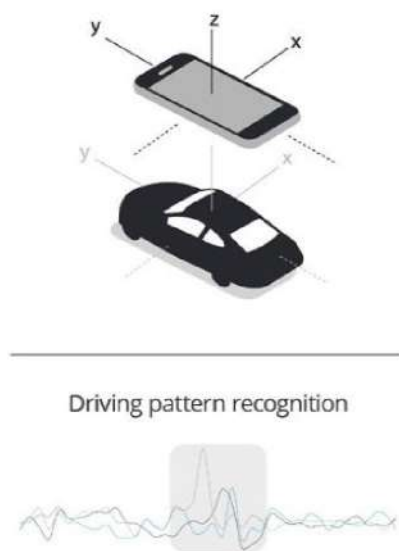
Αποθήκευση δεδομένων, θέματα ασφαλείας και ιδιωτικότητας

Τα δεδομένα αποθηκεύονται στο σύστημα **OSeven backend** χρησιμοποιώντας προηγμένες **τεχνικές κρυπτογράφησης και ασφάλειας δεδομένων**, σύμφωνα με τους εθνικούς νόμους και τις οδηγίες της ΕΕ για την προστασία των προσωπικών δεδομένων. Οι εφαρμογές που χρησιμοποιούνται υποστηρίζουν τον έλεγχο ταυτότητας του χρήστη και την κρυπτογράφηση για να αποτρέπεται η πρόσβαση σε μη εξουσιοδοτημένα δεδομένα από τρίτους.

Επεξεργασία δεδομένων

Αφού τα δεδομένα αποθηκευτούν στον εξυπηρετητή για να πραγματοποιηθεί η κεντρική επεξεργασία και η μείωση της διάστασης των δεδομένων, μετατρέπονται τα στοιχεία που έχουν συλλεχθεί σε σημαντικές συμπεριφορές και σε συναφείς με την οδική ασφάλεια παραμέτρους (δηλαδή χειρισμό και επεξεργασία μαζικών δεδομένων). Αυτό επιτυγχάνεται χρησιμοποιώντας τις

δύο μεγάλες μεθόδους επεξεργασίας δεδομένων που περιλαμβάνουν δύο οικογένειες τεχνικών, **τεχνικές εξόρυξης μεγάλων δεδομένων (data mining)** και αλγόριθμους **Machine Learning (ML)**.



Εικόνα4.3: Yaw, Pitch, Roll (Πηγή: OSeven)

Οι **μέθοδοι εκμάθησης μηχανών** (μέθοδοι φιλτραρίσματος, ομαδοποίησης και ταξινόμησης) χρησιμοποιούνται για τον καθαρισμό των δεδομένων από το θόρυβο, τα σφάλματα και την αναγνώριση επαναλαμβανόμενων μοτίβων μέσα στα δεδομένα. Στη συνέχεια, αυτά τα δεδομένα θα υποβληθούν σε επεξεργασία κατάλληλη για μεγάλο όγκο δεδομένων, προκειμένου να υπολογιστούν οι απαραίτητες παράμετροι και να ληφθούν δείκτες συμπεριφοράς που θα χρησιμοποιηθούν στην ανάλυση. Με άλλα λόγια, τα **δεδομένα χωρικά και χρονικά διαχωρισμένα** από το κινητό τηλέφωνο (Smartphone), επεξεργάζονται προκειμένου να προκύψουν **χρήσιμοι δείκτες οδικής ασφάλειας**. Οι μέθοδοι τεχνικής νοημοσύνης επιτρέπουν την **ανίχνευση της επιθετικής συμπεριφοράς** του οδηγού υπό τη μορφή απότομων συμβάντων, την απόσπαση προσοχής του οδηγού από τη χρήση του κινητού του τηλεφώνου, την αναγνώριση των τρόπων οδήγησης, την υπέρβαση του ορίου ταχύτητας, καθώς και ο προσδιορισμός των χρονικών και χωρικών χαρακτηριστικών όλων των παραπάνω.

Η διαδικασία των αλγορίθμων Machine Learning και η τεχνική εξόρυξης των Big Data δίνεται παρακάτω σε απλά βήματα:

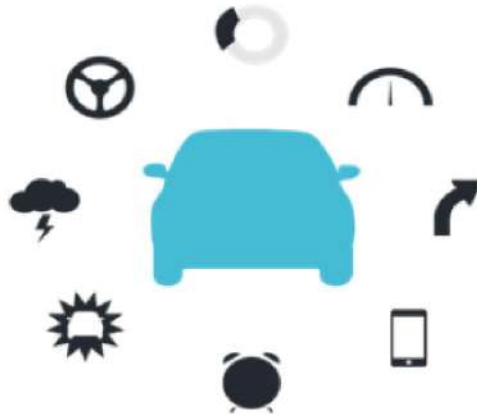
- Φιλτράρισμα δεδομένων και ανίχνευση τιμών που αποκλίνουν
- Εξομάλυνση δεδομένων (όπου χρειάζεται)
- Περιοχές υπερβολικής ταχύτητας
- Συμβάντα απότομων επιταχύνσεων
- Συμβάντα απότομων φρεναρισμάτων
- Συμβάντα απότομων στροφών
- Χρήση κινητού τηλεφώνου (ομιλία, σύνταξη μηνυμάτων, πλοήγηση στο διαδίκτυο)
- Οδήγηση εντός επικίνδυνων ωρών (απόσταση σε επικίνδυνες ώρες)
- Τρόπος μετακίνησης (αυτοκίνητο Ι.Χ., μέσα μαζικής μεταφοράς, ποδήλατο, μοτοποδήλατο)
- Καθορισμός οδηγού ή επιβάτη

Υπολογίζεται μια ποικιλία διαφορετικών μετά-δεδομένων έπειτα από τη διαδικασία Machine Learning (ML) τα οποία είναι χρήσιμα στον χρήστη ή στην αξιολόγηση της οδηγικής συμπεριφοράς, όπως για παράδειγμα:

Δείκτες έκθεσης κινδύνου

- Συνολική απόσταση (απόσταση σε μίλια)
- Διάρκεια οδήγησης
- Τύπος (τύποι) οδικού δικτύου που χρησιμοποιείται (η θέση δίνεται στο Παγκόσμιο Σύστημα Εντοπισμού θέσης (GPS) και ενσωματώνεται σε χορηγούς χαρτών, π.χ. Google, OSM)
- Περίοδος της ημέρας που πραγματοποιήθηκε η οδήγηση (ώρες αιχμής, επικίνδυνες ώρες)
- Καιρικές Συνθήκες
- Σκοπός του ταξιδιού

συνδυασμένα με άλλες πηγές δεδομένων (όρια ταχυτήτων και λεπτομερείς χάρτες ατυχημάτων).



Εικόνα4.4: Δείκτες κινδύνου οδήγησης (Πηγή: OSeven)

Δείκτες οδηγικής συμπεριφοράς

- Υπέρβαση του ορίου ταχύτητας (διάρκεια υπέρβασης του ορίου, υπέρβαση του ορίου ταχύτητας κλπ.)
- Αριθμός και σοβαρότητα απότομων συμβάντων
- Απότομο φρενάρισμα (επιτάχυνση κατά μήκος)
- Απότομη επιτάχυνση (επιτάχυνση κατά μήκος)
- Απότομες στροφές (γωνιακή ταχύτητα, πλευρική επιτάχυνση, πορεία)
- Οδηγική επιθετικότητα (π.χ. φρενάρισμα, επιτάχυνση)
- Απόσπαση λόγω χρήσης κινητού τηλεφώνου

Αυτοί οι δείκτες μαζί με άλλα δεδομένα (π.χ. δεδομένα από χάρτες) αξιοποιούνται διαδοχικά για να εφαρμοστεί η στατιστική ανάλυση ατομικά σε κάθε οδηγό, σε όλους τους τύπους οδών (αστική οδός, αυτοκινητόδρομος κλπ.) και κάτω από διάφορες συνθήκες οδήγησης, επιτρέποντας τη δημιουργία μιας μεγάλης βάσης δεδομένων με οδηγικά χαρακτηριστικά.

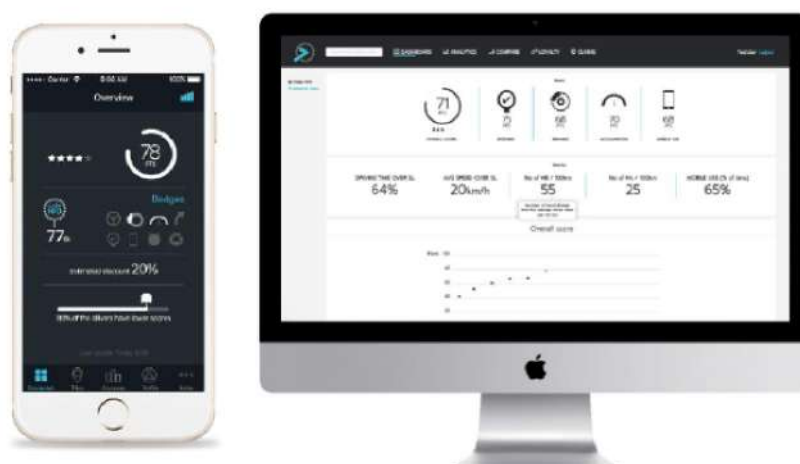
Το τελικό βήμα της διαδικασίας επεξεργασίας των δεδομένων είναι η ανάπτυξη του μοντέλου οδηγικής συμπεριφοράς. Συγκεντρωτικά δεδομένα

αναλύονται και το σύστημα αξιολόγησης βαθμονομείται με βάση το συνολικό δείγμα. Το μοντέλο οδηγικής συμπεριφοράς αναπτύσσει διάφορους δείκτες για την αξιολόγηση της οδήγησης και τελικώς συγκεντρώνει την όλη διαδικασία σε μια **βάση ανά ταξίδι-διαδρομή** για κάθε οδηγό στο δείγμα με σκοπό να παραχθεί το τελικό σύστημα βαθμολόγησης του κάθε οδηγού. Κάθε διαδρομή και επομένως κάθε οδηγός αξιολογείται με βάση τα χαρακτηριστικά που αναφέρθηκαν παραπάνω. Η τελική αξιολόγηση που παράγεται περιλαμβάνει τόσο συνολική αξιολόγηση όσο και αξιολόγηση ανά χαρακτηριστικό. Τα επεξεργασμένα δεδομένα μεταφέρονται σε **φιλικές προς το χρήστη εφαρμογές έξυπνων κινητών τηλεφώνων ή σε πλατφόρμες στο διαδίκτυο** που προσφέρουν φιλικά προς το χρήστη περιβάλλοντα για να λάβουν τις αναλυτικές αναφορές των διαδρομών τους. Η παρουσίαση της διαδικασίας των δεδομένων περιγράφεται στην επόμενη παράγραφο. Πρέπει να αναφερθεί ότι όλα τα επεξεργασμένα δεδομένα αποθηκεύονται στη βάση δεδομένων της OSeven ακολουθώντας διαδικασίες ανωνυμίας ώστε τα δεδομένα να μη μπορούν να συνδεθούν με συγκεκριμένο χρήστη.

Παρουσίαση δεδομένων

Τα αποτελέσματα όλης της προαναφερθείσας διαδικασίας είναι προσβάσιμα στην εφαρμογή έξυπνων κινητών τηλεφώνων (Smartphone) και στη Διαδικτυακή Πύλη, όπου είναι διαθέσιμα για το χρήστη να δει όλα τα συμβάντα που ανιχνεύθηκαν και τη θέση τους στο χάρτη όπως επίσης και όλα τα αποτελέσματα (συνολικά και ανά κατηγορία). Έτσι, προσφέρεται στον οδηγό ένας τρόπος φιλικός στον χρήστη να αντιλαμβάνεται τα τμήματα ταξιδιού με επικίνδυνη συμπεριφορά οδήγησης και να αποφύγει παρόμοιες συμπεριφορές στο μέλλον. Ταυτόχρονα, οι ασφαλιστικές εταιρείες έχουν πρόσβαση σε δεδομένα των πελατών τους χρησιμοποιώντας τη διαδικτυακή πύλη της OSeven. Τα αποτελέσματα οδήγησης χρησιμοποιούνται για τον καθορισμό των ασφάλιστρων ή/και προγραμμάτων αξιοπιστίας.

4. Συλλογή και Επεξεργασία Στοιχείων



Εικόνα4.5: Εφαρμογή στο κινητό και διαδικτυακή πύλη (Πηγή: OSeven)

Η βάση δεδομένων που προέκυψε δόθηκε σε αρχείο Microsoft Excel και περιλάμβανε τα μεγέθη / στοιχεία του Πίνακα 4.1.

Επιπλέον, προστέθηκαν στήλες με καινούργιες μεταβλητές οι οποίες κρίθηκαν απαραίτητες για την ανάλυση του στόχου της παρούσας Διπλωματικής. Συγκεκριμένα, δημιουργήθηκαν οι μεταβλητές του Πίνακα 4.2.

Πίνακας 4.1: Επεξήγηση Μεταβλητών Βάσης Δεδομένων

No	αύξων αριθμός
tripid	κωδική ονομασία διαδρομής
userid	κωδική ονομασία οδηγού
tripstarted	ημερομηνία/ώρα έναρξης διαδρομής
tripended	ημερομηνία/ώρα λήξης διαδρομής
duration	συνολική διάρκεια διαδρομής σε δευτερόλεπτα (s)
time_urban	συνολική διάρκεια οδήγησης σε αστικό περιβάλλον (s)
time_rural	συνολική διάρκεια οδήγησης σε υπεραστικό περιβάλλον (s)
time_rural/duration	ποσοστό χρόνου οδήγησης σε υπεραστικό περιβάλλον
time_highway	συνολική διάρκεια οδήγησης σε αυτοκινητόδρομο (s)
driving_duration	καθαρή διάρκεια οδήγησης χωρίς τον χρόνο για στάσεις σε δευτερόλεπτα (s)
driving_duration_urban	καθαρή διάρκεια οδήγησης σε αστικό περιβάλλον (s)
driving_duration_rural	καθαρή διάρκεια οδήγησης σε υπεραστικό περιβάλλον (s)
driving_duration_highwa	καθαρή διάρκεια οδήγησης σε αυτοκινητόδρομο (s)
totaldist	συνολική απόσταση διαδρομής σε χιλιόμετρα (km)

distance_urban	απόσταση διαδρομής σε αστικό περιβάλλον (km)
distance_rural	απόσταση διαδρομής σε υπεραστικό περιβάλλον (km)
distance_highway	απόσταση διαδρομής σε αυτοκινητόδρομο (km)
ha	απότομη επιτάχυνση (απόλυτος αριθμός)
hb	απότομο φρενάρισμα/επιβράδυνση (απόλυτος αριθμός)
hc	απότομη στροφή (απόλυτος αριθμός)
ha_urban	απότομη επιτάχυνση σε αστικό περιβάλλον
ha_rural	απότομη επιτάχυνση σε υπεραστικό περιβάλλον
ha_highway	απότομη επιτάχυνση σε αυτοκινητόδρομο
hb_urban	απότομη επιβράδυνση σε αστικό περιβάλλον
hb_rural	απότομη επιβράδυνση σε υπεραστικό περιβάλλον
hb_highway	απότομη επιβράδυνση σε αυτοκινητόδρομο
hc_urban	απότομη στροφή σε αστικό περιβάλλον
hc_rural	απότομη στροφή σε υπεραστικό περιβάλλον
hc_highway	απότομη στροφή σε αυτοκινητόδρομο
avaccel	μέση επιτάχυνση (km/h/s)

avdecel	μέση επιβράδυνση (km/h/s)
smooth_corner	μέση στροφική ταχύτητα (o/s)
av_accel_urban	μέση επιτάχυνση σε αστικό περιβάλλον (km/h/s)
av_accel_rural	μέση επιτάχυνση σε υπεραστικό περιβάλλον (km/h/s)
av_accel_highway	μέση επιτάχυνση σε αυτοκινητόδρομο (km/h/s)
av_decel_urban	μέση επιβράδυνση σε αστικό περιβάλλον (km/h/s)
av_decel_rural	μέση επιβράδυνση σε υπεραστικό περιβάλλον (km/h/s)
av_decel_highway	μέση επιβράδυνση σε αυτοκινητόδρομο (km/h/s)
mobileUsage	ποσοστό της καθαρής διάρκειας οδήγησης (driving_duration) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο
mobileUsage_urban	ποσοστό της καθαρής διάρκειας οδήγησης σε αστικό περιβάλλον (driving_duration_urban) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο
mobileUsage_rural	ποσοστό της καθαρής διάρκειας οδήγησης σε υπεραστικό περιβάλλον (driving_duration_rural) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο
mobileUsage_highway	ποσοστό της καθαρής διάρκειας οδήγησης σε αυτοκινητόδρομο (driving_duration_highway) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο
perc_speeding	ποσοστό της καθαρής διάρκειας οδήγησης (driving_duration) με ταχύτητα πάνω από το όριο ταχύτητας

perc_speeding_urban	ποσοστό της καθαρής διάρκειας οδήγησης σε αστικό περιβάλλον (driving_duration_urban) με ταχύτητα πάνω από το όριο ταχύτητας
perc_speeding_rural	ποσοστό της καθαρής διάρκειας οδήγησης σε υπεραστικό περιβάλλον (driving_duration_rural) με ταχύτητα πάνω από το όριο ταχύτητας
perc_speeding_highway	ποσοστό της καθαρής διάρκειας οδήγησης σε αυτοκινητόδρομο (driving_duration_highway) με ταχύτητα πάνω από το όριο ταχύτητας
av_speeding	μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας
av_speeding_urban	μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε αστικό περιβάλλον
av_speeding_rural	μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε υπεραστικό περιβάλλον
av_speeding_highway	μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε αυτοκινητόδρομο
avgspeed	μέση ταχύτητα (km/h)
maxspeed	μέγιστη τιμή ταχύτητας (km/h)
speedq25	τιμή που αντιστοιχεί στο 25% της καταγεγραμμένης ταχύτητας
speedq50	τιμή που αντιστοιχεί στο 50% της καταγεγραμμένης ταχύτητας
speedq75	τιμή που αντιστοιχεί στο 75% της καταγεγραμμένης ταχύτητας
StSpDev	τυπική απόκλιση ταχύτητας (km/h)
avjerk	μέσος όρος ρυθμού μεταβολής της επιτάχυνσης (km/h/s ²)

avjerkp	θετικός ρυθμός μεταβολής της επιτάχυνσης (km/h/s ²)
avjerkn	μειούμενος ρυθμός μεταβολής της επιτάχυνσης (km/h/s ²)
riskyhoursdistance	απόσταση οδήγησης από τις 22:00 μέχρι τις 05:00

Πίνακας 4.2: Επεξήγηση Νέων Μεταβλητών που προστέθηκαν στη Βάση Δεδομένων

timeMobileUsage	χρόνος σε λεπτά (min) που χρησιμοποιήθηκε το κινητό τηλέφωνο από τον χρήστη κατά την καθαρή διάρκεια οδήγησης
sum_speeding	χρόνος σε λεπτά (min) κατά την καθαρή διάρκεια οδήγησης με ταχύτητα πάνω από το όριο
cs_ha	σωρευτικό άθροισμα γεγονότων απότομων επιταχύνσεων (απόλυτος αριθμός)
cs_hb	σωρευτικό άθροισμα γεγονότων απότομων φρεναρισμάτων (απόλυτος αριθμός)
cs_TMU	σωρευτικό άθροισμα του χρόνου χρήσης κινητού τηλεφώνου κατά την καθαρή διάρκεια οδήγησης σε λεπτά (min)
cs_sum_speeding	σωρευτικό άθροισμα του χρόνου υπέρβασης του ορίου ταχύτητας σε λεπτά (min)
cs_totaldist	σωρευτικό άθροισμα της συνολικής απόστασης διαδρομής σε χιλιόμετρα (km)
cs_duration	σωρευτικό άθροισμα της καθαρής διάρκειας οδήγησης χωρίς τον χρόνο για στάσεις σε λεπτά (min)
cs_ha / cs_totaldist	σωρευτικό άθροισμα γεγονότων απότομων επιταχύνσεων προς σωρευτικό άθροισμα της συνολικά διανυθείσας απόστασης (απόλυτος αριθμός / km)
cs_hb / cs_totaldist	σωρευτικό άθροισμα γεγονότων απότομων φρεναρισμάτων προς σωρευτικό άθροισμα της συνολικά διανυθείσας απόστασης (απόλυτος αριθμός / km)

cs_TMU / cs_duration	σωρευτικό άθροισμα του χρόνου χρήσης κινητού τηλεφώνου κατά την καθαρή διάρκεια οδήγησης προς σωρευτικό άθροισμα καθαρής διάρκειας οδήγησης (min / min, ως ποσοστό δηλαδή)
cs_sum_speeding/cs_duration	σωρευτικό άθροισμα του χρόνου υπέρβασης του ορίου ταχύτητας κατά την καθαρή διάρκεια οδήγησης προς σωρευτικό άθροισμα καθαρής διάρκειας οδήγησης (min / min, ως ποσοστό δηλαδή)

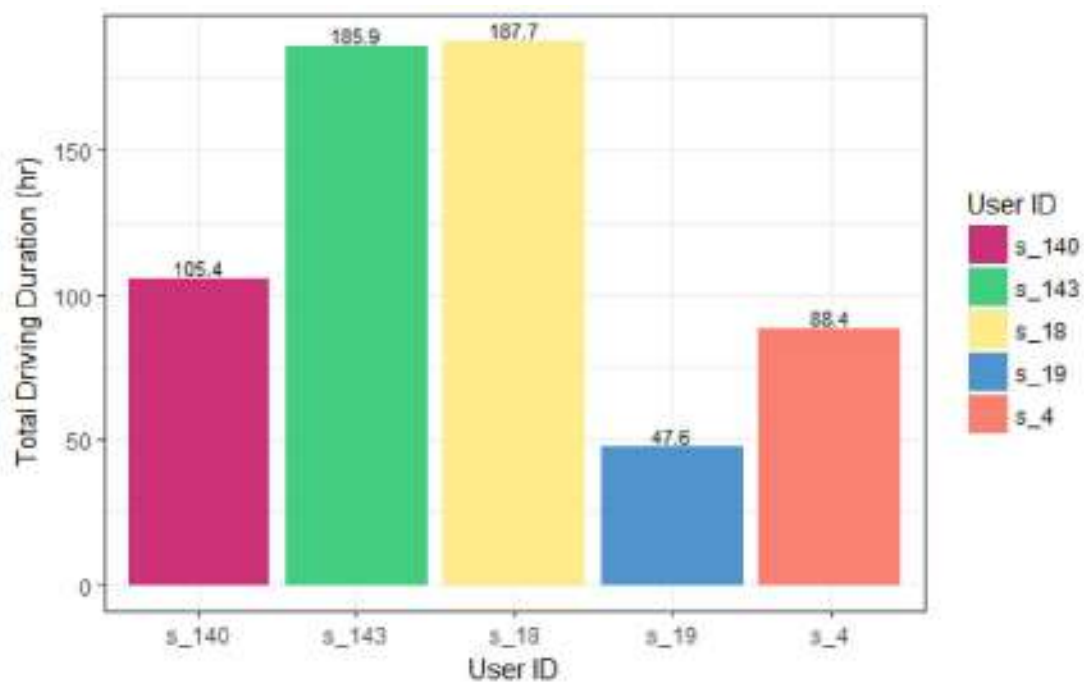
4.3 Προκαταρκτική Ανάλυση Στοιχείων

Τα Διαγράμματα αποτελούν μια προκαταρκτική ανάλυση, η οποία συμβάλλει στην βαθύτερη κατανόηση της βάσης δεδομένων και θα χρησιμοποιηθεί για την εξαγωγή ποιοτικότερων συμπερασμάτων. Για την προκαταρκτική αυτή ανάλυση δεν χρησιμοποιήθηκαν όλοι οι οδηγοί, αλλά έγινε μια ενδεικτική επιλογή **5 οδηγών** από τη βάση δεδομένων, για λόγους ευκρίνειας των απεικονίσεων στα παρακάτω διαγράμματα. Οι οδηγοί αυτοί επιλέχθηκαν με βάση το κριτήριο της διαφορετικότητας των **χαρακτηριστικών της συνολικής οδηγικής συμπεριφοράς μεταξύ τους** (π.χ. επιθετικότητα, συχνότητα χρήσης κινητού τηλεφώνου κατά την οδήγηση, αριθμός και διάρκεια ταξιδιών κ.α.), ούτως ώστε να γίνει ευκολότερα κατανοητή η ανάγκη του προσδιορισμού του χρόνου παρατήρησης κάθε οδηγού.

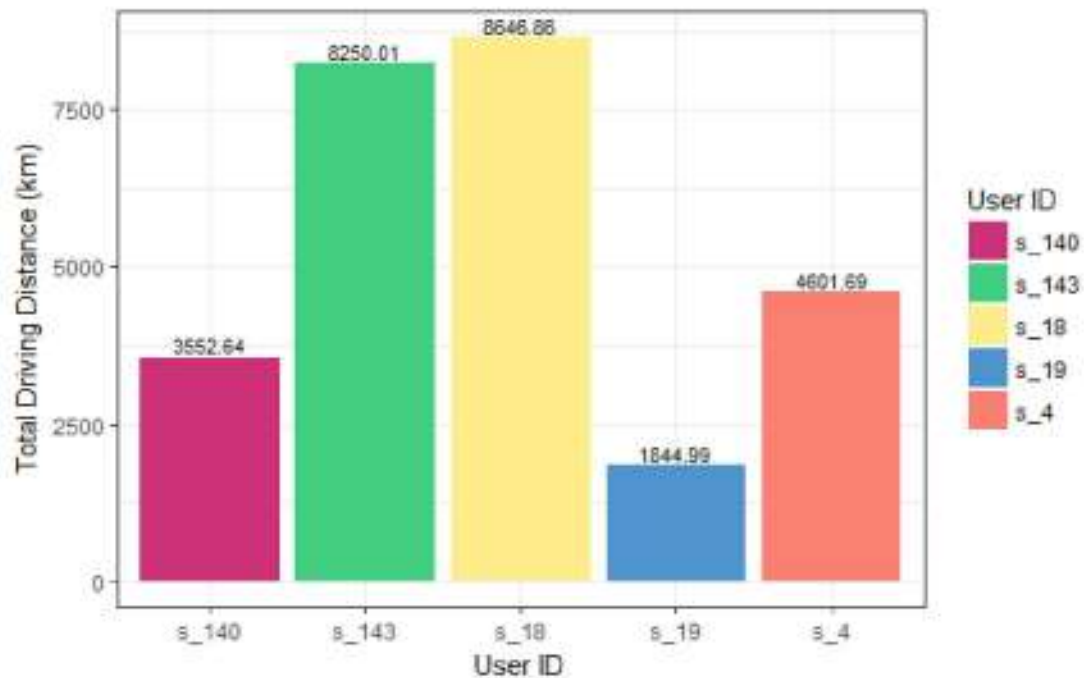
Αξίζει να αναφερθεί ότι για την φάση της προκαταρκτικής ανάλυσης που ακολουθεί, η βάση δεδομένων ήταν **ταξινομημένη ανά οδηγό** και με τα ταξίδια κάθε οδηγού σε **χρονολογική σειρά** (από το παλαιότερο στο πιο πρόσφατο). Αυτό έχει σημασία, διότι σε κάποια από τα διαγράμματα που ακολουθούν ο οριζόντιος άξονας αφορά στην σωρευτικά αθροιζόμενη διάρκεια ή απόσταση των ταξιδιών μέχρι εκείνη τη χρονική στιγμή, **χωρίς αυτό να σημαίνει ότι τα ταξίδια είχαν ομοιογενή χαρακτηριστικά** (π.χ. παρόμοια διάρκεια ή παρόμοια συνολική απόσταση).

Τα Διαγράμματα 4.1 και 4.2 δείχνουν αντίστοιχα ότι η Συνολική Καθαρή Διάρκεια Οδήγησης, καθώς και τα Συνολικά Διανυόμενα Χιλιόμετρα (κατ' αντιστοιχία) διαφέρουν σημαντικά μεταξύ των παραπάνω οδηγών, με κάποιους να φτάνουν τις 190 ώρες και τα 8650 km οδήγησης κατά τη διάρκεια του πειράματος, και άλλους να μην ξεπερνούν τις 50 ώρες και τα 1900 km αντίστοιχα.

4. Συλλογή και Επεξεργασία Στοιχείων



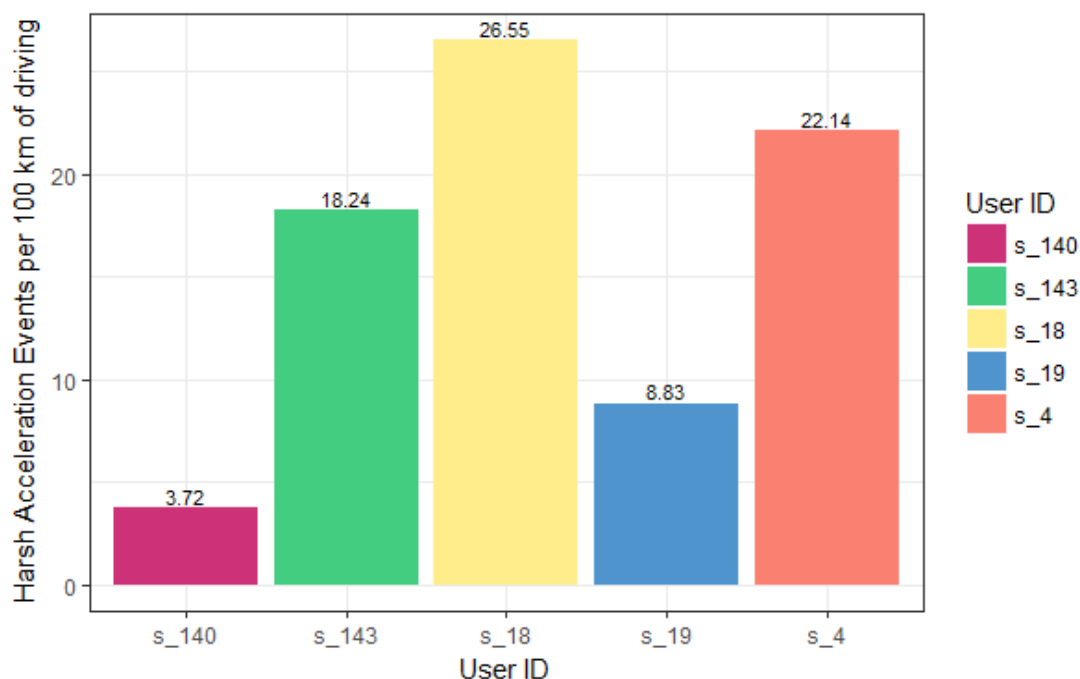
Διάγραμμα4.1: Συνολική Καθαρή Διάρκεια Οδήγησης για κάθε οδηγό (σε ώρες)



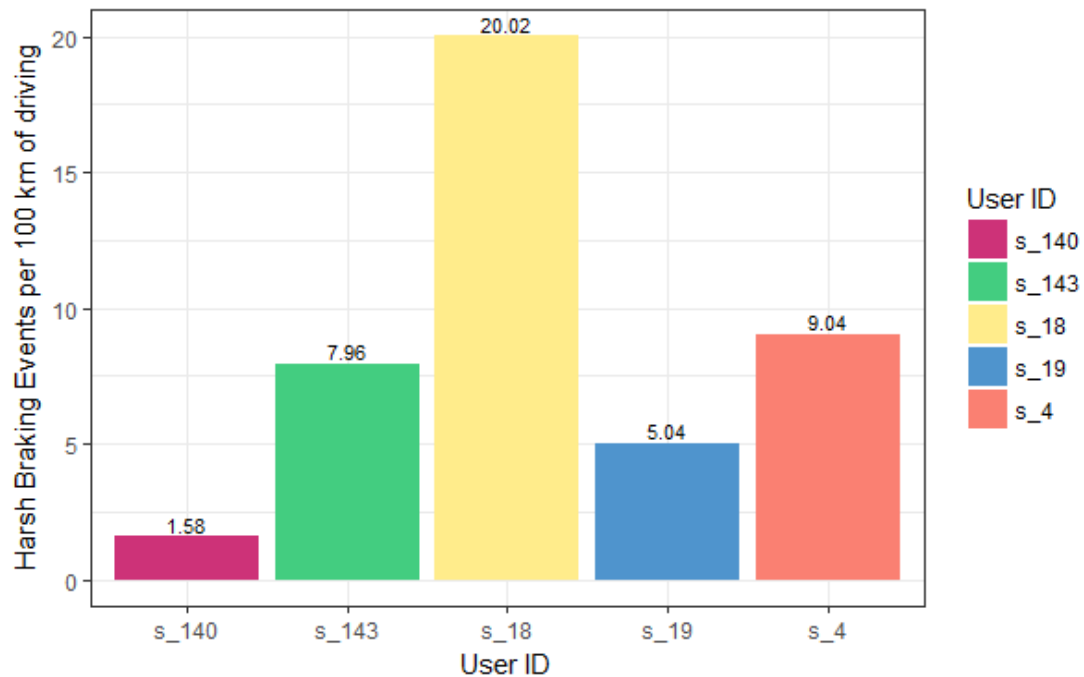
Διάγραμμα4.2: Συνολική Διανυθείσα Απόσταση για κάθε οδηγό (σε χιλιόμετρα)

Παρατηρείται ότι η γενική εικόνα των Διαγραμμάτων 4.3 και 4.4 είναι όμοια, δηλαδή οι ίδιοι οδηγοί που παρουσιάζουν υψηλές τιμές $ha/100km$ παρουσιάζουν και υψηλές τιμές $hb/100km$, ενώ αντίστοιχα, οι ίδιοι που παρουσιάζουν χαμηλές τιμές $ha/100km$ παρουσιάζουν και χαμηλές τιμές $hb/100km$. Αυτό γενικά μας δίνει την δυνατότητα να διακρίνουμε τους πλέον επιθετικούς οδηγούς από τους πλέον προσεκτικούς. Ο αριθμός των γεγονότων Απότομων Επιταχύνσεων, ανηγμένος ανά 100 km οδήγησης διαφέρει σημαντικά από οδηγό σε οδηγό, με τους πιο επιθετικούς να ξεπερνούν τα 20 $ha/100km$ και τους πιο προσεκτικούς να μην ξεπερνούν τα 5 $ha/100km$.

Ο αριθμός των γεγονότων Απότομων Φρεναρισμάτων, ανηγμένος ανά 100 km οδήγησης διαφέρει επίσης σημαντικά από οδηγό σε οδηγό, με τους πιο επιθετικούς να βρίσκονται στα 20 $hb/100km$ και τους πιο προσεκτικούς να μην ξεπερνούν τα 2 $hb/100km$.



Διάγραμμα 4.3: Αριθμός γεγονότων απότομων επιταχύνσεων ανά 100 km οδήγησης για κάθε οδηγό

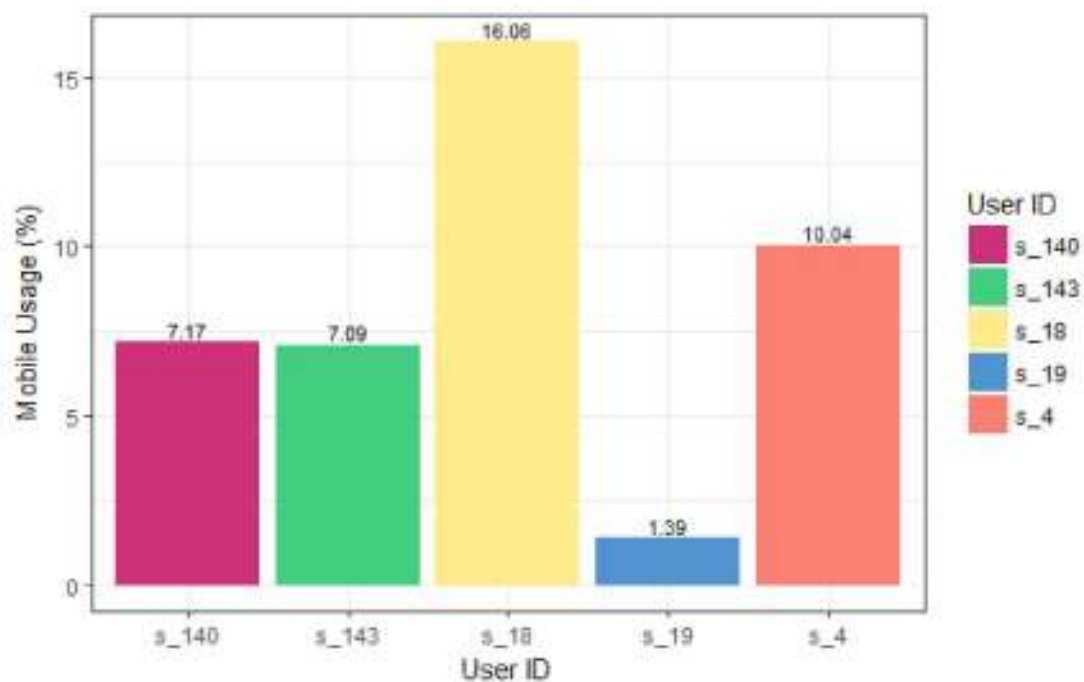


Διάγραμμα4.4: Αριθμός γεγονότων απότομων επιβραδύνσεων ανά 100 km οδήγησης για κάθε οδηγό

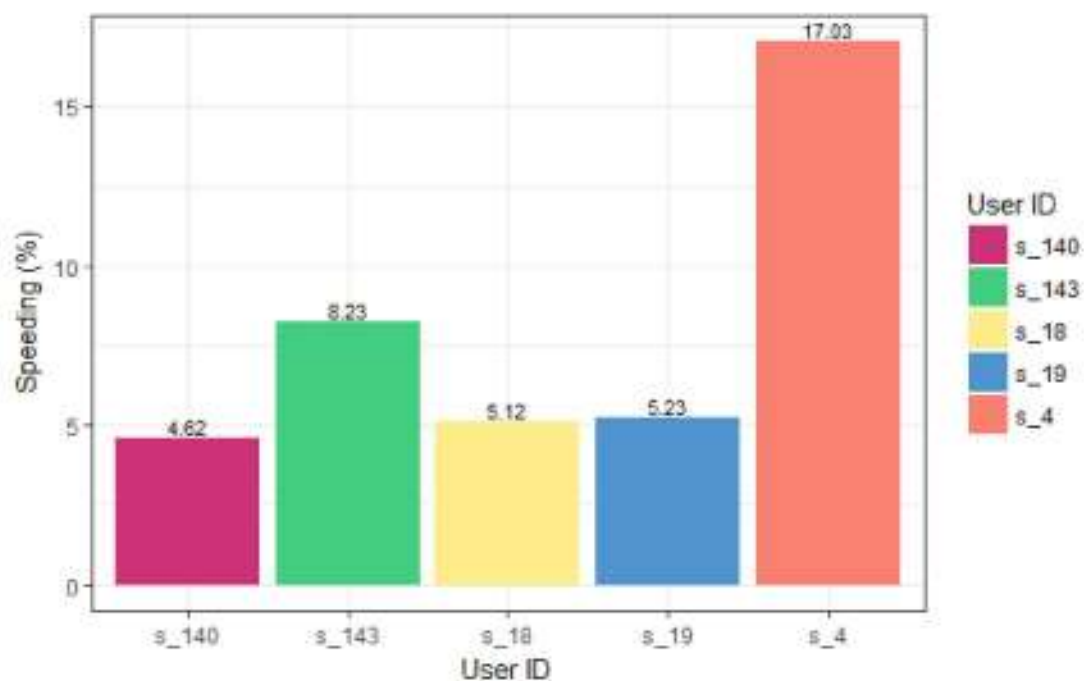
Όπως φαίνεται στα Διαγράμματα 4.5 και 4.6, το ποσοστό (%) του Συνολικού Καθαρού Χρόνου Οδήγησης κατά το οποίο χρησιμοποιήθηκε το κινητό τηλέφωνο από τους οδηγούς παρουσιάζει σημαντική διαφοροποίηση με τρεις από τους πέντε οδηγούς να το χρησιμοποιούν 7-10% του Συνολικού Καθαρού Χρόνου Οδήγησης. Αξίζει επίσης να αναφερθεί ότι η μεγαλύτερη τιμή εμφανίζεται στον οδηγό με τους μεγαλύτερους αριθμούς ha/100km και hb/100km (16%).

Το ποσοστό (%) του Συνολικού Καθαρού Χρόνου Οδήγησης κατά το οποίο υπήρξε υπέρβαση του ορίου ταχύτητας κυμαίνεται στα 5-8% για τους παραπάνω οδηγούς, ενώ μόνο ένας φαίνεται να ξεπερνάει το 17%.

4. Συλλογή και Επεξεργασία Στοιχείων

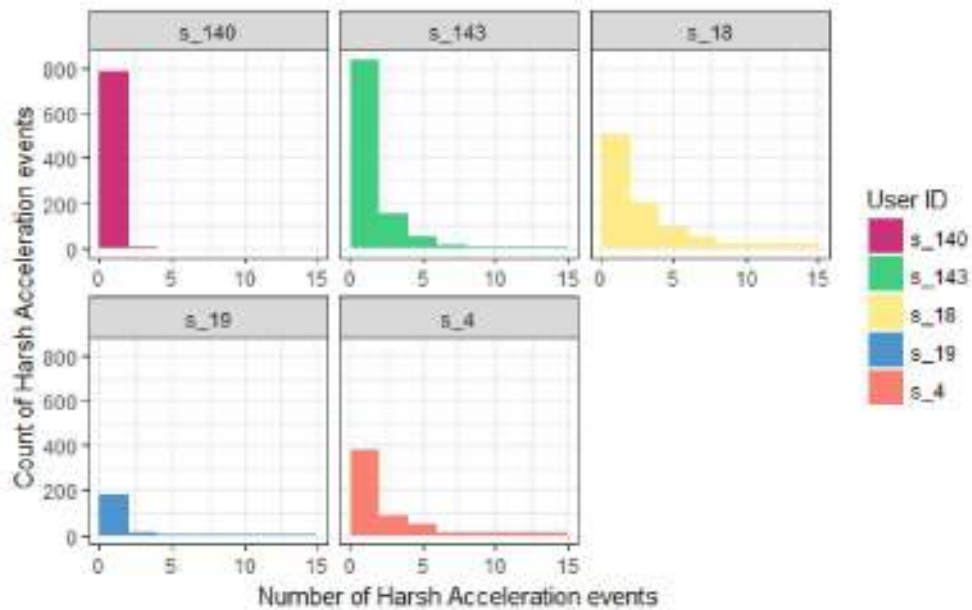


Διάγραμμα4.5: Πόσοστό (%) του Συνολικού Καθαρού Χρόνου Οδήγησης κατά το οποίο χρησιμοποιήθηκε το κινητό τηλέφωνο, για κάθε οδηγό

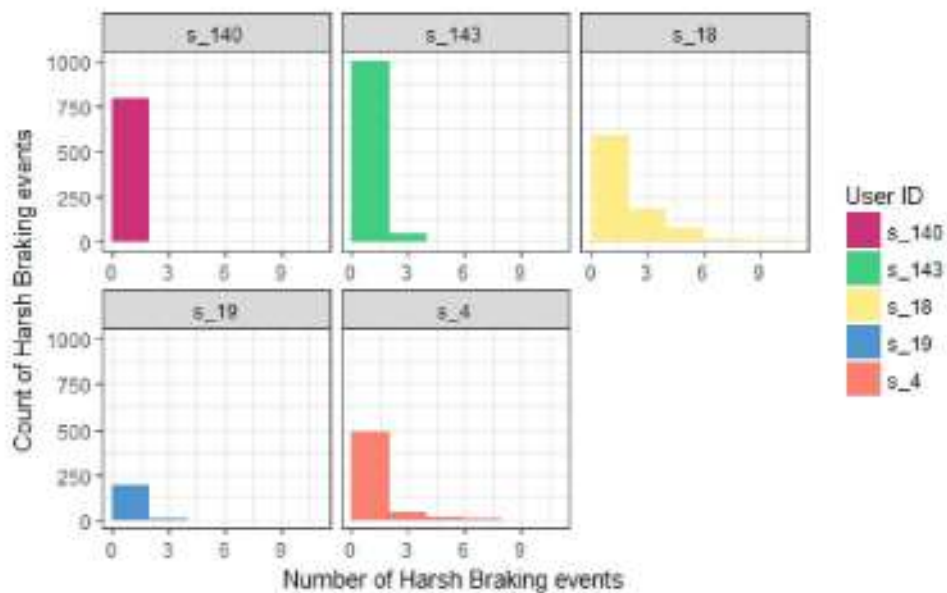


Διάγραμμα4.6: Ποσοστό (%) του Συνολικού Καθαρού Χρόνου οδήγησης κατά το οποίο υπήρξε υπέρβαση του ορίου ταχύτητας, για κάθε οδηγό

Από τα Ιστογράμματα Συχνοτήτων των Απότομων Επιταχύνσεων και των Απότομων Φρεναρισμάτων των Διαγραμμάτων 4.7 και 4.8, προκύπτει ότι οι οδηγοί που γενικά είχαν πιο επιθετική συμπεριφορά, έχουν και πιο «απλωμένα» ιστογράμματα συχνοτήτων, ενώ αυτά των πιο προσεκτικών οδηγών είναι συγκεντρωμένα γύρω από τις τιμές 0-2 γεγονότων h_a ή h_b .

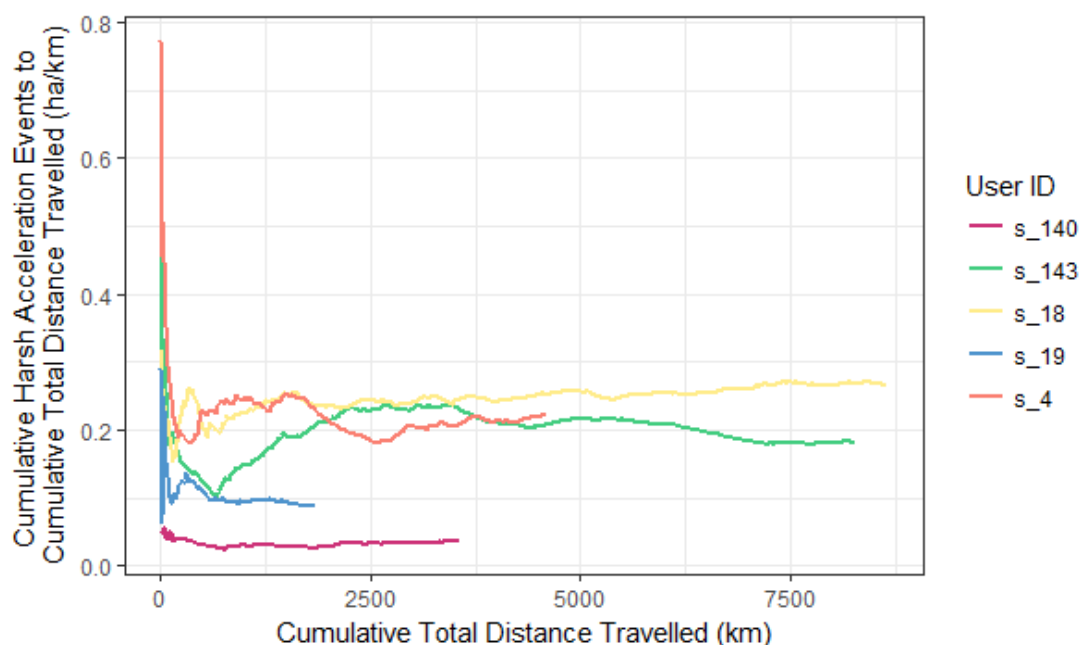


Διάγραμμα 4.7: Ιστόγραμμα Συχνοτήτων του αριθμού των Απότομων Επιταχύνσεων για κάθε οδηγό



Διάγραμμα 4.8: Ιστόγραμμα Συχνοτήτων του αριθμού των Απότομων Φρεναρισμάτων για κάθε οδηγό

Συνεχίζοντας την προκαταρκτική ανάλυση, και για να γίνει πιο εύκολα αντιληπτή και να οπτικοποιηθεί η μεταβολή των παραπάνω οδηγικών χαρακτηριστικών των οδηγών, δημιουργήθηκαν τα Διαγράμματα 4.9 – 4.12 των νέων μεταβλητών σωρευτικών αθροισμάτων.



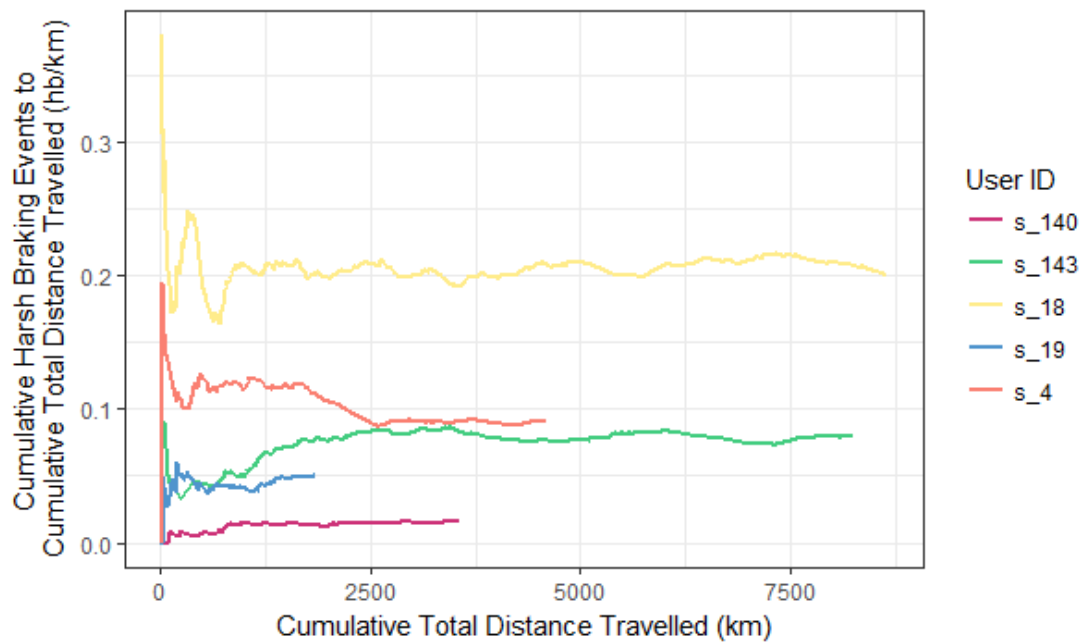
Διάγραμμα 4.9: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων Απότομων Επιταχύνσεων προς τα σωρευτικά αθροιζόμενα km οδήγησης (κατακόρυφος άξονας) ως προς τη συνολική απόσταση οδήγησης (km) για κάθε οδηγό

Στο Διάγραμμα 4.9 απεικονίζεται η μεταβολή του μεγέθους:

$$\frac{\text{Cumulative Sum (Harsh Acceleration Events)}}{\text{Cumulative Sum (Total Distance Travelled)}}$$

(τιμές στον άξονα y) με την πάροδο των διανυθέντων χιλιομέτρων (τιμές στον άξονα x). Είναι εμφανές ότι τα συνολικά χιλιόμετρα οδήγησης για κάθε χρήστη είναι διαφορετικά, γι' αυτό και δεν εξετάστηκε ο αριθμός των ha σαν απόλυτος αριθμός αλλά έγινε η αναγωγή του στην απόσταση που έχει διανύσει κάθε οδηγός ώστε να είναι συγκρίσιμα (ha/km). Είναι επίσης εμφανές ότι κάθε χρήστης έχει διαφορετικό αριθμό ha/km στον οποίο συγκλίνει η συμπεριφορά

του, με τον μικρότερο να είναι αυτός του χρήστη s_140 και ο μεγαλύτερος αυτός των s_4 και s_18.

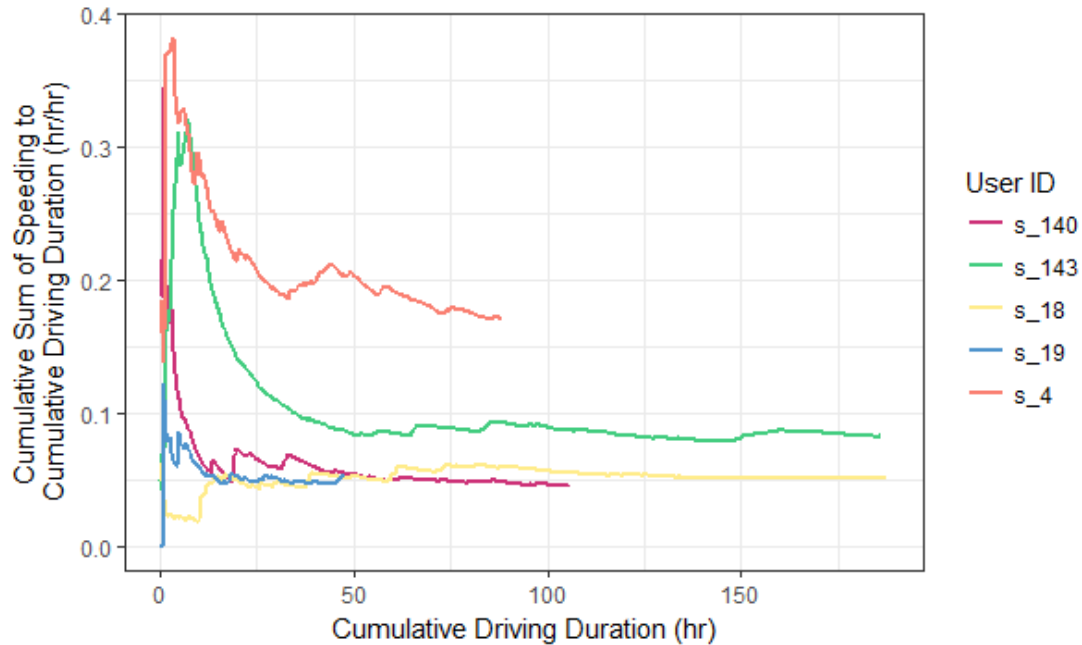


Διάγραμμα 4.10: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων Απότομων Φρεναρισμάτων προς τα σωρευτικά αθροιζόμενα km οδήγησης (κατακόρυφος άξονας) ως προς τη συνολική απόσταση οδήγησης (km) για κάθε οδηγό

Αντίστοιχα, στο Διάγραμμα 4.10 απεικονίζεται η μεταβολή του μεγέθους:

$$\frac{\text{Cumulative Sum (Harsh Braking Events)}}{\text{Cumulative Sum (Total Distance Travelled)}}$$

(τιμές στον άξονα y) με την πάροδο των διανυθέντων χιλιομέτρων (τιμές στον άξονα x). Είναι εμφανές ότι τα συνολικά χιλιόμετρα οδήγησης για κάθε χρήστη είναι διαφορετικά, γι' αυτό και δεν εξετάστηκε ο αριθμός των hb σαν απόλυτος αριθμός αλλά έγινε η αναγωγή του στην απόσταση που έχει διανύσει κάθε οδηγός ώστε να είναι συγκρίσιμα (hb/km). Είναι επίσης εμφανές ότι κάθε χρήστης έχει διαφορετικό αριθμό hb/km στον οποίο συγκλίνει η συμπεριφορά του, με τον μικρότερο να είναι αυτός του χρήστη s_140 και ο μεγαλύτερος αυτός των s_4 και s_18.

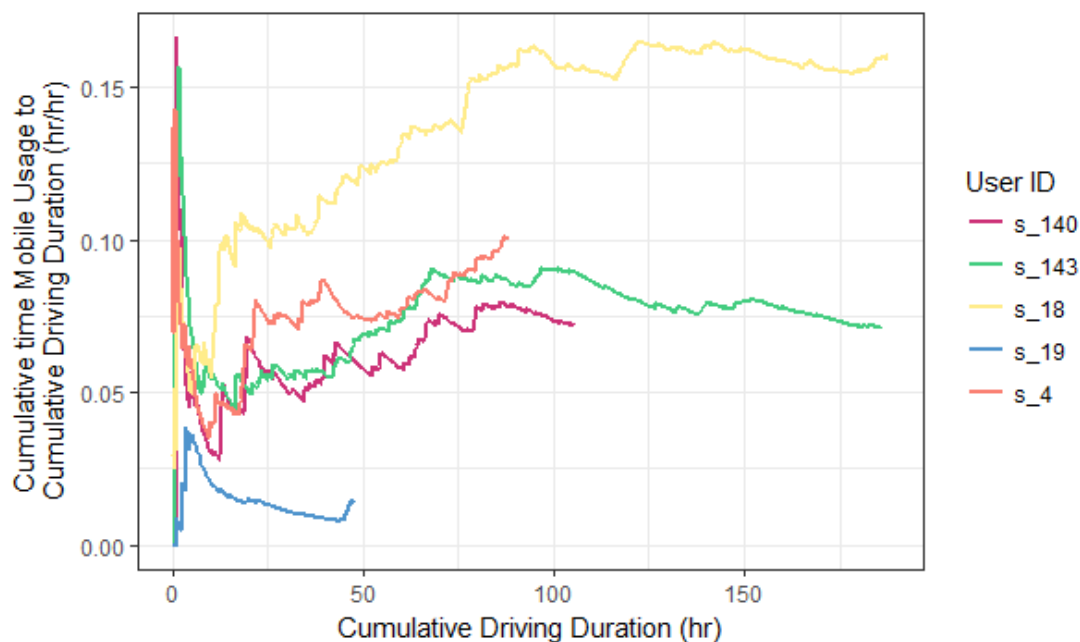


Διάγραμμα 4.11: Συγκετρωτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων ωρών υπέρβασης του ορίου ταχύτητας προς τη σωρευτικά αθροιζόμενη διάρκεια οδήγησης (κατακόρυφος άξονας – hr/hr(%) ως προς τη συνολική διάρκεια οδήγησης (hr) για κάθε οδηγό

Στο Διάγραμμα 4.11 απεικονίζεται η μεταβολή του μεγέθους:

$$\frac{\text{Cumulative Sum (Time Speeding)}}{\text{Cumulative Sum (Driving Duration)}}$$

(τιμές στον άξονα y) με την συνολική διάρκεια οδήγησης (τιμές στον άξονα x). Είναι εμφανές ότι η διάρκεια οδήγησης για κάθε χρήστη είναι διαφορετική, γι' αυτό και δεν εξετάστηκε ο αριθμός των ωρών υπέρβασης του ορίου ταχύτητας σαν απόλυτος αριθμός αλλά έγινε η αναγωγή του στην διάρκεια οδήγησης. Είναι επίσης εμφανές ότι κάθε χρήστης έχει διαφορετικό αριθμό ωρών υπέρβασης του ορίου ταχύτητας στον οποίο συγκλίνει η συμπεριφορά του, με τον μικρότερο να είναι αυτός του χρήστη s_140 και ο μεγαλύτερος αυτός του s_4.



Διάγραμμα 4.12: Συγκριτικό - Συγκριτικό διάγραμμα απεικόνισης της μεταβολής του μεγέθους των σωρευτικά αθροιζόμενων ωρών χρήσης κινητού τηλεφώνου προς τη σωρευτικά αθροιζόμενη διάρκεια οδήγησης (κατακόρυφος άξονας – hr/hr(%)) ως προς τη συνολική διάρκεια οδήγησης (hr) για κάθε οδηγό

Τέλος, στο Διάγραμμα 4.12 απεικονίζεται η μεταβολή του μεγέθους:

$$\frac{\text{Cumulative Sum (Time Mobile Usage)}}{\text{Cumulative Sum (Driving Duration)}}$$

(τιμές στον άξονα y) με την συνολική διάρκεια οδήγησης (τιμές στον άξονα x). Είναι εμφανές ότι η διάρκεια οδήγησης για κάθε χρήστη είναι διαφορετική, γι' αυτό και δεν εξετάστηκε ο αριθμός των MU(hr) σαν απόλυτος αριθμός αλλά έγινε η αναγωγή του στην διάρκεια οδήγησης. Είναι επίσης εμφανές ότι, σε αντίθεση με τα προηγούμενα, δεν είναι ξεκάθαρα εμφανής η σύγκλιση του κάθε χρήστη σε κάποιο αριθμό MU(hr)/hr.

Από την παραπάνω προκαταρκτική ανάλυση προκύπτει ότι:

- **Οι πιο προσεκτικοί οδηγοί** συγκλίνουν σε **μικρότερους αριθμούς ha/km και hb/km** από τους **πιο επιθετικούς**.

- Το διάγραμμα του χρόνου υπέρβασης του ορίου ταχύτητας φαίνεται να παρουσιάζει **κοινό χρονικό σημείο σύγκλισης** για τους 5 οδηγούς γύρω στις **50 ώρες** (Διάγραμμα 4.11).
- Το μέγεθος του **χρόνου χρήσης κινητού τηλεφώνου** δεν φαίνεται να **συγκλίνει τόσο ξεκάθαρα** όσο τα υπόλοιπα χαρακτηριστικά που εξετάζουμε.
- Οι πλέον επιθετικοί οδηγοί τείνουν να παρουσιάζουν **συγκλίνουσα συμπεριφορά** πιο «γρήγορα» από τους πιο προσεκτικούς οδηγούς.
- Η **ποσοτικοποίηση του απαιτούμενου χρόνου παρατήρησης** ενός οδηγού φαίνεται να **εξαρτάται** τόσο από την **επιθετικότητά** του, όσο και από το **διαφορετικό χαρακτηριστικό οδηγικής συμπεριφοράς** που μελετάται κάθε φορά.

5. ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Εισαγωγή

Στο παρόν κεφάλαιο περιγράφονται αναλυτικά τα βήματα που ακολουθήθηκαν κατά την εφαρμογή της μεθοδολογίας και παρουσιάζονται τα αποτελέσματα των διαδικασιών. Αρχικά, δημιουργούνται οι υποβάσεις δεδομένων ανάλογα με τη διάρκεια και ανάλογα με τον οδηγό, ταυτόχρονα με τη δημιουργία των νέων μεταβλητών των μεγεθών σωρευτικών αθροισμάτων (CumulativeSums). Έτσι, προκύπτουν οι μεταβλητές οι οποίες θα χρησιμοποιηθούν στην ανάλυση, οι οποίες είναι **cs_ha / cs_totaldist**, **cs_hb / cs_totaldist**, **cs_TMU / cs_duration**, **cs_sum_speeding / cs_duration**, και των οποίων η σύγκλιση θα εξεταστεί. Στη συνέχεια, εξετάζεται η μεταβλητότητα των παραπάνω μεγεθών. Για το σκοπό αυτό δημιουργούνται τα μεγέθη του Κινητού Μέσου Όρου (MovingAverage) και της Μεταβλητότητας (Volatility) και καταρτίζονται τα επιθυμητά όρια και οι συνθήκες που θα πρέπει να ικανοποιούνται, έτσι ώστε να μπορεί να θεωρηθεί ότι μια μεταβλητή συγκλίνει σε κάποια τιμή y ύστερα από t χρονικό διάστημα παρατήρησης. Από τα αποτελέσματα αυτών των τιμών σύγκλισης για κάθε οδηγό και για κάθε διάρκεια, εξάγονται τα συμπεράσματα της διπλωματικής εργασίας.

5.2 Προετοιμασία των Δεδομένων

Όπως έχει αναφερθεί και στο Κεφάλαιο 3, η βάση δεδομένων που χρησιμοποιήθηκε αποτελείται από **21.610** ξεχωριστά **ταξίδια-διαδρομές 68 οδηγών**, και επομένως ήταν απαραίτητη η δημιουργία ενός τύπου «ομαδοποίησης» και χρονολογικής ταξινόμησης των δεδομένων προκειμένου να γίνει εφικτή η παρακολούθηση της μεταβολής των μεγεθών των χαρακτηριστικών οδηγικής συμπεριφοράς με την πάροδο του χρόνου. Η εισαγωγή της βάσης δεδομένων στην RStudio γίνεται με τη μορφή υπολογιστικού φύλλου με τιμές διαχωρισμένες με κόμμα (.csv) και η ομαδοποίηση πραγματοποιήθηκε με τη σύνταξη κώδικα, με τη βοήθεια του

οποίου δημιουργήθηκαν οι «υποβάσεις» δεδομένων, ανάλογα με **την καθαρή διάρκεια οδήγησης των ταξιδιών**.

Πρωτού πραγματοποιηθεί ο χωρισμός της βάσης δεδομένων στις παραπάνω υποβάσεις όμως, είχε προηγηθεί η δημιουργία των μεταβλητών των μεγεθών των οποίων εξετάζεται η συμπεριφορά, δηλαδή των **cs_ha / cs_totaldist, cs_hb / cs_totaldist, cs_TMU / cs_duration, cs_sum_speeding / cs_duration**. Για τη δημιουργία των μεταβλητών αυτών συντάχθηκε επίσης κώδικας σε R.

Στη συνέχεια, πραγματοποιήθηκε μια ακόμα «ομαδοποίηση» στις υποβάσεις, αυτή τη φορά ως προς **τον οδηγό**. Και για αυτή την δεύτερη ομαδοποίηση, όπως και για την πρώτη, έγινε χρήση ειδικού κώδικα που συντάχθηκε στην R. Έτσι, δημιουργήθηκαν οι τελικές υποβάσεις που χρησιμοποιήθηκαν ως δεδομένα εισόδου στα Υπολογιστικά Φύλλα Excel που χρησιμοποιήθηκαν στη συνέχεια των αναλύσεων.

5.3 Διερεύνηση Μεταβλητότητας – Σύγκλισης των χαρακτηριστικών οδηγικής συμπεριφοράς

Για την διερεύνηση της μεταβλητότητας και της σύγκλισης των παραπάνω χαρακτηριστικών, χρησιμοποιήθηκαν διάφορες μέθοδοι, των οποίων η επιλογή επεξηγήθηκε στο Κεφάλαιο 3. Προκειμένου όμως να υλοποιηθούν οι παραπάνω μέθοδοι, δημιουργήθηκαν Υπολογιστικά Φύλλα Excel, τα οποία όπως προαναφέρθηκε, χρησιμοποιούν ως δεδομένα εισόδου τις υποβάσεις δεδομένων που δημιουργήθηκαν μέσω της R.

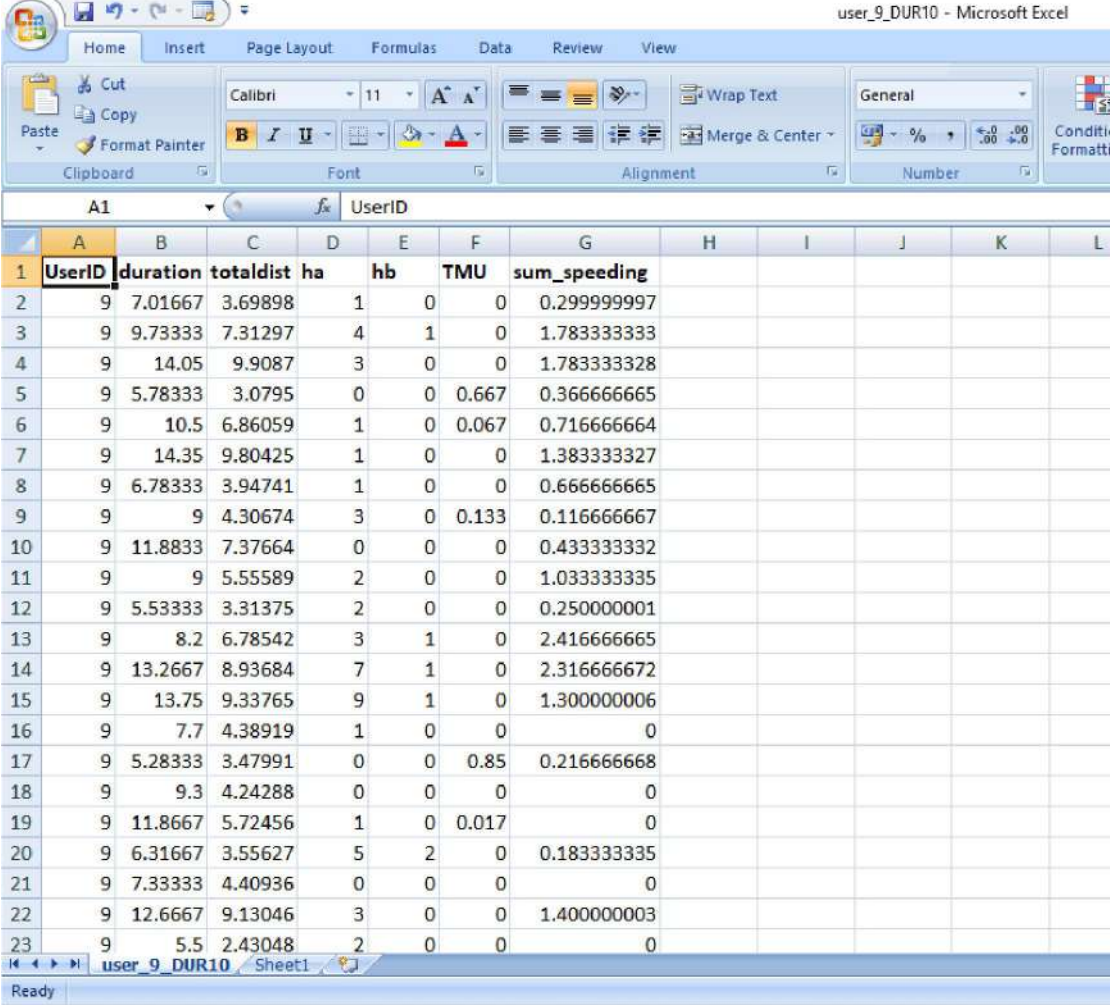
5.3.1 Περιγραφή της εισαγωγής των δεδομένων εισόδου στα Υπολογιστικά Φύλλα Excel

Στην παρακάτω εικόνα απεικονίζεται η πρώτη καρτέλλα (με τίτλο **“RInputData”**) του αρχείου **“RESULTS”**, το οποίο δημιουργήθηκε στο Microsoft Excel. Στην καρτέλλα αυτή, εισάγονται τα δεδομένα εισόδου από τις υποβάσεις δεδομένων που δημιουργήθηκαν μέσω της R, στις **στήλες B έως H**, όπως φαίνεται παρακάτω.

	A	B	C	D	E	F	G	H	I	J	K	L
1	number of trips	UserID	duration	totaldist	ha	hb	TMU	sum_speeding		CumSum Metrics	cs_duration	cs_totald
2		1	9	7.01667	3.69898	1	0	0	0.299999997		7.0166667	3.69898
3		2	9	9.73333	7.31297	4	1	0	1.783333333		16.75	11.0119
4		3	9	14.05	9.9087	3	0	0	1.783333328		30.8	20.9206
5		4	9	5.78333	3.0795	0	0	0.67	0.366666665		36.583333	24.0001
6		5	9	10.5	6.86059	1	0	0.07	0.716666664		47.083333	30.8607
7		6	9	14.35	9.80425	1	0	0	1.383333327		61.433333	40.6650
8		7	9	6.78333	3.94741	1	0	0	0.666666665		68.216667	44.6124
9		8	9	9	4.30674	3	0	0.13	0.116666667		77.216667	48.9191
10		9	9	11.8833	7.37664	0	0	0	0.433333332		89.1	56.2957
11		10	9	9	5.55589	2	0	0	1.033333335		98.1	61.8516
12		11	9	5.53333	3.31375	2	0	0	0.250000001		103.63333	65.1654
13		12	9	8.2	6.78542	3	1	0	2.416666665		111.83333	71.9508
14		13	9	13.2667	8.93684	7	1	0	2.316666672		125.1	80.8876

Εικόνα 5.1: Πεδίο εισαγωγής δεδομένων εισόδου στο αρχείο **“RESULTS”**

Οι υποβάσεις δεδομένων έχουν τη μορφή που φαίνεται στην παρακάτω εικόνα (Εικόνα 5.2). Κάνοντας αντιγραφή των δεδομένων από την υποβάση, και στη συνέχεια επικόλληση στο πεδίο εισαγωγής της καρτέλλας **“RInputData”** του αρχείου **“RESULTS”**, όπως φαίνεται στην παραπάνω εικόνα (Εικόνα 5.1), υπολογίζονται αυτόματα (μέσω συναρτήσεων) τα υπόλοιπα μεγέθη που χρησιμοποιήθηκαν στην ανάλυση.



user_9_DUR10 - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J	K	L
1	UserID	duration	totaldist	ha	hb	TMU	sum_speeding					
2	9	7.01667	3.69898	1	0	0	0.299999997					
3	9	9.73333	7.31297	4	1	0	1.783333333					
4	9	14.05	9.9087	3	0	0	1.783333328					
5	9	5.78333	3.0795	0	0	0.667	0.366666665					
6	9	10.5	6.86059	1	0	0.067	0.716666664					
7	9	14.35	9.80425	1	0	0	1.383333327					
8	9	6.78333	3.94741	1	0	0	0.666666665					
9	9	9	4.30674	3	0	0.133	0.116666667					
10	9	11.8833	7.37664	0	0	0	0.433333332					
11	9	9	5.55589	2	0	0	1.033333335					
12	9	5.53333	3.31375	2	0	0	0.250000001					
13	9	8.2	6.78542	3	1	0	2.416666665					
14	9	13.2667	8.93684	7	1	0	2.316666672					
15	9	13.75	9.33765	9	1	0	1.300000006					
16	9	7.7	4.38919	1	0	0	0					
17	9	5.28333	3.47991	0	0	0.85	0.216666668					
18	9	9.3	4.24288	0	0	0	0					
19	9	11.8667	5.72456	1	0	0.017	0					
20	9	6.31667	3.55627	5	2	0	0.183333335					
21	9	7.33333	4.40936	0	0	0	0					
22	9	12.6667	9.13046	3	0	0	1.400000003					
23	9	5.5	2.43048	2	0	0	0					

Εικόνα 5.2: Παράδειγμα μορφής δεδομένων εισόδου, όπως λαμβάνονται από τις υποβάσεις

Στην ίδια καρτέλλα, λίγο δεξιότερα, υπολογίζονται αυτόματα (με χρήση συναρτήσεων) τα **μεγέθη σωρευτικών αθροισμάτων (CumSumMetrics)**, καθώς επίσης και η **μεταβλητότητα (Volatility)** των μεγεθών αυτών, χρησιμοποιώντας χρονικό παράθυρο 20 ταξιδιών. Η αναλυτική επεξήγηση των συναρτήσεων που χρησιμοποιήθηκαν, όπως επίσης και η αντίστοιχη αιτιολόγηση για την επιλογή τους, έχει προαναφερθεί στο Κεφάλαιο 3 που αφορά τη Μεθοδολογική Προσέγγιση.

5.3.2 Περιγραφή της εξαγωγής των δεδομένων εξόδου από τα Υπολογιστικά Φύλλα Excel

Στο αρχείο “RESULTS”, έχουν επίσης διαμορφωθεί οι καρτέλλες: “**HA**”, “**HB**”, “**TMU**”, “**SUM_SPEEDING**” και “**OUTPUT**”. Οι καρτέλλες αυτές

παίρνουν τα δεδομένα που χρησιμοποιούν από την καρτέλλα **“RInputData”**. Στη συνέχεια, οι καρτέλλες: **“HA”**, **“HB”**, **“TMU”** και **“SUM_SPEEDING”** υπολογίζουν, τον ελάχιστο αριθμό των ταξιδιών διάρκειας *ήπου πρέπει να παρατηρηθεί ένας οδηγός, ικανοποιώντας τα κριτήρια που έχουν τεθεί για την παραδοχή σύγκλισης συμπεριφοράς, για **συμβάντα απότομης επιτάχυνσης, συμβάντα απότομης πέδησης**, απόσπαση προσοχής λόγω **χρήσης κινητού τηλεφώνου** και οδήγησης με **υπέρβαση του ορίου ταχύτητας**, αντίστοιχα. Στην καρτέλλα **“OUTPUT”** αποθηκεύονται προσωρινά οι παραπάνω τιμές σύγκλισης για κάθε μεταβλητή, και επομένως, αποτελούν τα δεδομένα εξόδου της ανάλυσης, τα οποία αντιγράφονται και μεταφέρονται στον συγκεντρωτικό πίνακα **“CONVERGENCE”**, ενός νέου αρχείου Excel, όπου και αποθηκεύονται.*

Παρακάτω περιγράφεται πιο αναλυτικά, η διαδικασία που ακολουθείται σε κάθε καρτέλλα προκειμένου να προκύψει ο ελάχιστος αριθμός ταξιδιών διάρκειας *ήπου πρέπει να παρατηρηθεί ένας οδηγός, ικανοποιώντας τα κριτήρια που έχουν τεθεί για την παραδοχή σύγκλισης συμπεριφοράς. Για να γίνει πιο εύκολα κατανοητό, χρησιμοποιείται ως παράδειγμα η καρτέλλα “HA”, αλλά η διαδικασία που ακολουθείται είναι ακριβώς αντίστοιχη και για τις άλλες μεταβλητές.*

Στην καρτέλλα **“HA”** υπάρχουν οι στήλες **UserID**, **duration**, **ha**, **cs_ha**, **totaldist**, **cs_totaldist**, **cs_ha / cs_totaldist** και **volatility** οι οποίες παίρνουν τις τιμές τους από την καρτέλλα **“RInputData”**. Η καρτέλλα χωρίζεται σε δύο τμήματα: το πρώτο αφορά στην ανάλυση του μεγέθους **cs_ha / cs_totaldist** και στην διερεύνηση σύγκλισής του, και το δεύτερο αφορά στην **ανάλυση της μεταβλητότητας του μεγέθους cs_ha / cs_totaldist**, δηλαδή στην διερεύνηση σύγκλισης του **volatility(cs_ha / cs_totaldist)**. Τα δύο τμήματα απεικονίζονται στις παρακάτω Εικόνες 5.3 και 5.4. Επιπροσθέτως, στην καρτέλλα υπάρχει ένα πεδίο από το οποίο ορίζονται οι παράμετροι των ελέγχων που διενεργούνται και οι τιμές των ορίων (Εικόνα 5.5).

5. Ανάλυση και Αποτελέσματα

	G	H	I	J	K	L	M	N	O	P
	cs_ha / cs_totaldist	Moving Average of (cs_ha / cs_totaldist)	% change	CONDITION: LB(cs_ha/cs_totaldist) < UB	5 Consecutive trips with (% change) < ORIO	5 Consecutive trips with % change of cs_ha / cs_totaldist within the limitations	CS Upper Bound	CS Lower Bound	CONVERGENCE (CumSum)	# TRIP CONV. (CumSum)
1										
2	0.27034444	#N/A								
3	0.4540518	#N/A	67.95% ekstos				0	0.448904927	0.397505083	0
4	0.382397038	#N/A	15.78% ekstos				0	0.448904927	0.397505083	0
5	0.33333107	#N/A	12.83% ekstos				0	0.448904927	0.397505083	0
6	0.291632521	#N/A	12.51% ekstos				0	0.448904927	0.397505083	0
7	0.245911668	#N/A	15.68% ekstos		ekstos change %		0	0.448904927	0.397505083	0
8	0.246568105	#N/A	0.27% ekstos		ekstos change %		0	0.448904927	0.397505083	0
9	0.286186442	#N/A	16.07% ekstos		ekstos change %		0	0.448904927	0.397505083	0
10	0.248686409	#N/A	13.10% ekstos		ekstos change %		0	0.448904927	0.397505083	0
11	0.258683337	0.301779283	4.02% ekstos		ekstos change %		0	0.448904927	0.397505083	0
12	0.27622007	0.302366846	6.78% ekstos		ekstos change %		0	0.448904927	0.397505083	0
13	0.291865892	0.286148255	5.66% ekstos		ekstos change %		0	0.448904927	0.397505083	0
14	0.346158955	0.282524447	18.60% ekstos		ekstos change %		0	0.448904927	0.397505083	0
15	0.410084321	0.290199772	18.47% ekstos		ekstos change %		0	0.448904927	0.397505083	0
16	0.401629632	0.301199483	2.06% ekstos		ekstos change %		0	0.448904927	0.397505083	0
17	0.38738177	0.315345493	3.55% ekstos		ekstos change %		0	0.448904927	0.397505083	0
18	0.371321036	0.327821786	4.15% ekstos		ekstos change %		0	0.448904927	0.397505083	0
19	0.360904322	0.335293574	2.81% ekstos		ekstos change %		0	0.448904927	0.397505083	0
20	0.39420112	0.349845046	9.23% ekstos		ekstos change %		0	0.448904927	0.397505083	0
21	0.379220404	0.361898752	3.80% ekstos		ekstos change %		0	0.448904927	0.397505083	0

Εικόνα 5.3: Καρτέλλα "HA" - Τμήμα ελέγχων για τη σύγκλιση του $cs_ha / cs_totaldist$

	Q	R	S	T	U	V	W	X	Y	Z
	Volatility (tw: 20)	Moving Average of Volatility	% change	CONDITION: V_LB < (Volatility) < V_UB	5 Consecutive trips with (% change) < ORIO	5 Consecutive trips with Volatility % change within the limitations	Volatility Upper Bound	Volatility Lower Bound	CONVERGENCE (VOLATILITY)	# TRIP CONV. (volatility)
1										
23	0.046388503	#N/A	32.88% ekstos				0	0.009258274	0	0
24	0.043353683	#N/A	6.54% ekstos				0	0.009258274	0	0
25	0.041157245	#N/A	5.07% ekstos				0	0.009258274	0	0
26	0.038697845	#N/A	5.98% ekstos				0	0.009258274	0	0
27	0.034307817	#N/A	11.34% ekstos				0	0.009258274	0	0
28	0.034827778	#N/A	1.52% ekstos		ekstos change %		0	0.009258274	0	0
29	0.032422327	#N/A	6.91% ekstos		ekstos change %		0	0.009258274	0	0
30	0.028335795	#N/A	12.60% ekstos		ekstos change %		0	0.009258274	0	0
31	0.028727319	0.03973279	1.38% ekstos		ekstos change %		0	0.009258274	0	0
32	0.02840568	0.035662399	1.12% ekstos		ekstos change %		0	0.009258274	0	0
33	0.028197036	0.035843253	0.73% ekstos		ekstos change %		0	0.009258274	0	0
34	0.023245756	0.03183246	17.56% ekstos		ekstos change %		0	0.009258274	0	0
35	0.015987709	0.029315506	31.22% ekstos		ekstos change %		0	0.009258274	0	0
36	0.015909693	0.027036691	0.49% ekstos		ekstos change %		0	0.009258274	0	0
37	0.015604455	0.025166355	1.92% ekstos		ekstos change %		0	0.009258274	0	0
38	0.015145592	0.023198136	2.94% ekstos		ekstos change %		0	0.009258274	0	0
39	0.014905921	0.021446496	1.58% ekstos		ekstos change %		0	0.009258274	0	0
40	0.013229978	0.019935914	11.24% ekstos		ekstos change %		0	0.009258274	0	0
41	0.012805221	0.018343704	3.21% ekstos		ekstos change %		0	0.009258274	0	0

Εικόνα 5.4: Καρτέλλα "HA" - Τμήμα ελέγχων για τη σύγκλιση του μεγέθους μεταβλητότητας του $cs_ha / cs_totaldist$ (Volatility[$cs_ha / cs_totaldist$])

AA	AB
OPIA METRIC	OPIA VOLATILITY
OPIO % CHANGE	OPIO % CHANGE
1.50%	1.50%
1	1
ANQ OPIO CumSum	ANQ OPIO VOLATILITY
0.448904927	0.009258274
KATQ OPIO CumSum	KATQ OPIO VOLATILITY
0.397505083	-0.003260519

Εικόνα 5.5: Πεδίο εισαγωγής τιμών ορίων και παραμέτρων ελέγχων

Στο πεδίο αυτό καθορίζεται:

- η τιμή ορίου της ποσοστιαίας μεταβολής του κινητού μέσου όρου των **cs_ha / cs_totaldist** και **volatility(cs_ha / cs_totaldist)**, αντίστοιχα. Και στις δύο περιπτώσεις έχουμε επιλέξει το όριο να είναι 1.50% (0.015).
- Η τιμή του παράγοντα **k** που πολλαπλασιάζει την τυπική απόκλιση στον υπολογισμό των άνω και κάτω ορίων, μέσω της σχέσης: *Μέση Τιμή* $\pm$ **k** * *Τυπική Απόκλιση*. Και στις δύο περιπτώσεις έχουμε επιλέξει η τιμή του **k** να είναι ίση με τη μονάδα.
- Υπολογίζονται τα Άνω και Κάτω Όρια των **cs_ha / cs_totaldist** και **volatility(cs_ha / cs_totaldist)**, μέσω της προηγούμενης σχέσης.

Αφού έχουν οριστεί οι παραπάνω τιμές, εξετάζεται αν και πότε ικανοποιούνται οι συνθήκες που βρίσκονται στις υπόλοιπες στήλες. Η στήλη του Κινητού Μέσου Όρου (MovingAverage) υπολογίζει τον κινητό μέσο όρο της στήλης **cs_ha / cs_totaldist** ή **volatility(cs_ha / cs_totaldist)** αντίστοιχα, χρησιμοποιώντας κινητό παράθυρο 20 παρατηρήσεων. Στη συνέχεια, η στήλη “% change” υπολογίζει την ποσοστιαία μεταβολή (σε απόλυτη τιμή) μεταξύ διαδοχικών τιμών της στήλης του Κινητού Μέσου Όρου.

Η στήλη – συνθήκη “**CONDITION: LB < (cs_ha / cs_totaldist) < UB**” ελέγχει εάν η τιμή του Κινητού Μέσου Όρου του μεγέθους **cs_ha / cs_totaldist** βρίσκεται εντός των άνω και κάτω ορίων που υπολογίστηκαν από την σχέση: *Μέση Τιμή* $\pm$ **k** * *Τυπική Απόκλιση*, παίρνοντας την τιμή «entos» αν η συνθήκη είναι αληθής, ή την τιμή «ektos» αν η συνθήκη είναι ψευδής. Αντίστοιχα και για τη στήλη “**CONDITION: V_LB < (Volatility) < V_UB**”, η οποία αναφέρεται στα αντίστοιχα μεγέθη και στήλες για το μέγεθος του **volatility(cs_ha / cs_totaldist)**.

Η στήλη “**5 Consecutive trips with (% change) < ORIO**” ελέγχει αν υπάρχουν 5 διαδοχικά ταξίδια, για τα οποία ισχύει ότι η ποσοστιαία μεταβολή (σε απόλυτη τιμή) μεταξύ διαδοχικών τιμών της στήλης του Κινητού Μέσου Όρου είναι μικρότερη ή ίση με το όριο που έχει τεθεί για τις ποσοστιαίες

μεταβολές. Αν η συνθήκη είναι αληθής, το αντίστοιχο κελί παίρνει την τιμή «entoschange %», διαφορετικά παίρνει την τιμή «ektoschange %».

Η στήλη **“5 Consecutive trips with % change of cs_ha / cs_totaldist within the limitations”** βρίσκει το πρώτο ταξίδι για το οποίο η προηγούμενη στήλη (“5 Consecutive trips with (% change) <ORIO”) παίρνει την τιμή «entos change %». Μόλις εντοπίσει το κελί αυτό, παίρνει σαν τιμή την μέση τιμή των παρατηρήσεων **cs_ha / cs_totaldist** (αυτή η επιλογή έγινε κυρίως για λόγους οπτικοποίησης στα διαγράμματα, καθώς αν χρησιμοποιούνταν οι τιμές 0 και 1 υπήρχαν προβλήματα με τις κλίμακες των αξόνων των διαγραμμάτων). Αντίστοιχα και για τη στήλη **“5 Consecutive trips with Volatility % change within the limitations”**, η οποία αναφέρεται στα αντίστοιχα μεγέθη και στήλες για το μέγεθος του **volatility(cs_ha / cs_totaldist)**.

Τέλος, η στήλη **“CONVERGENCE (CumSum)”** και αντίστοιχα η στήλη **“CONVERGENCE (Volatility)”**, ελέγχουν αν ικανοποιούνται ταυτόχρονα όλες οι παρακάτω συνθήκες:

- Ο Κινητός Μέσος Όρος βρίσκεται εντός των ορίων (η στήλη “CONDITION: $LB < (cs_ha / cs_totaldist) < UB$ ” ή αντίστοιχα η “CONDITION: $V_LB < (Volatility) < V_UB$ ” παίρνουν στο συγκεκριμένο κελί την τιμή «entos»).
- Έχουν βρεθεί 5 διαδοχικά ταξίδια για τα οποία ισχύει ότι η ποσοστιαία μεταβολή (σε απόλυτη τιμή) μεταξύ διαδοχικών τιμών της στήλης του Κινητού Μέσου Όρου είναι μικρότερη ή ίση με το όριο που έχει τεθεί για τις ποσοστιαίες μεταβολές (δηλαδή η στήλη “5 Consecutive trips with (% change) < ORIO” έχει λάβει την τιμή «entoschange %»).
- Η τιμή του Κινητού Μέσου Όρου στο αντίστοιχο κελί αποτελεί τοπικό ακρότατο (το κριτήριο αυτό εξασφαλίζει ότι οι γειτονικές τιμές του Κινητού Μέσου Όρου είναι μικρότερες ή μεγαλύτερες του σημείου που έχει επιλεγεί, και επομένως δεν ανήκει σε κάποια ακολουθία σημείων τα οποία έχουν μια συγκεκριμένη τάση π.χ. αύξουσα ή φθίνουσα).

Όταν βρεθεί το κελί για το οποίο ικανοποιούνται όλες οι παραπάνω συνθήκες, το αντίστοιχο κελί της στήλης **“CONVERGENCE (CumSum)”**(ή αντίστοιχα

της “**CONVERGENCE (Volatility)**”) παίρνει την τιμή «OK», που σημαίνει ότι έχει βρεθεί το σημείο από το οποίο και μετά το συγκεκριμένο χαρακτηριστικό συγκλίνει σε μια συγκεκριμένη τιμή. Στην δίπλα στήλη “**# TRIP CONV. (CumSum)**” ή αντίστοιχα στην “**# TRIPCONV. (Volatility)**”, καταγράφεται ο αύξων αριθμός του ταξιδιού αυτού. Για παράδειγμα, μπορούμε να πούμε ότι το μέγεθος **cs_ha / cs_totaldist** για ταξίδια μέσης διάρκειας 10 λεπτών του οδηγού με ονομασία χρήστη «9» συγκλίνει μετά από 58 ταξίδια, ενώ το μέγεθος του **volatility(cs_ha / cs_totaldist)** συγκλίνει μετά από 169 ταξίδια.

Η διαδικασία αυτή, όπως αναφέρθηκε και προηγουμένως είναι ακριβώς η ίδια και για τις υπόλοιπες μεταβλητές που εξετάζονται, δηλαδή τις: **cs_hb / cs_totaldist**, **cs_TMU / cs_duration**, **cs_sum_speeding / cs_duration** και τα αντίστοιχα μεγέθη **Volatility** αυτών.

Τελικώς, στην καρτέλλα “**OUTPUT**” αποθηκεύονται προσωρινά οι παραπάνω χρόνοι σύγκλισης για κάθε μεταβλητή, τόσο για το ίδιο το μέγεθος, όσο και για τη μεταβλητότητά του, τα οποία αποτελούν τα δεδομένα εξόδου της ανάλυσης. Τα δεδομένα αυτά στη συνέχεια αντιγράφονται και μεταφέρονται στον συγκεντρωτικό πίνακα “**CONVERGENCE**”, ενός νέου αρχείου Excel, όπου και αποθηκεύονται. Ταυτόχρονα αποθηκεύονται και οι μέσες τιμές των χαρακτηριστικών CumSum και της μεταβλητότητάς τους, οι οποίες αφορούν και τις τιμές σύγκλισης των μεταβλητών αυτών.

5.4 Παρουσίαση Αποτελεσμάτων Ανάλυσης

Η παραπάνω διαδικασία εφαρμόστηκε για τους 68 οδηγούς για τις υποβάσεις δεδομένων που αφορούσαν σε ταξίδια μέσης διάρκειας μικρότερης των 5 λεπτών, 10 λεπτών και 20 λεπτών. Η ανάλυση πραγματοποιήθηκε για τις παραπάνω μέσες διάρκειες καθότι ο αριθμός των ταξιδιών με διάρκεια μεγαλύτερη των 25 λεπτών ήταν σημαντικά μικρότερος, με αποτέλεσμα αρκετοί οδηγοί να έχουν λιγότερα από 30 ταξίδια στις υποβάσεις αυτές, ή και πολλές φορές κανένα απολύτως. Και σε αυτόν τον χωρισμό όμως,

παρατηρήθηκαν αντίστοιχα προβλήματα, καθώς αρκετοί οδηγοί δεν είχαν γενικά ικανοποιητικό αριθμό παρατηρήσεων, και κατά συνέπεια δεν χρησιμοποιήθηκαν στην ανάλυση, αφού δεν θα παρουσίαζαν καμία τάση ή σύγκλιση στην οδηγική τους συμπεριφορά. Τελικά χρησιμοποιήθηκαν τα δεδομένα από **29 οδηγούς** για την εξαγωγή των αποτελεσμάτων.

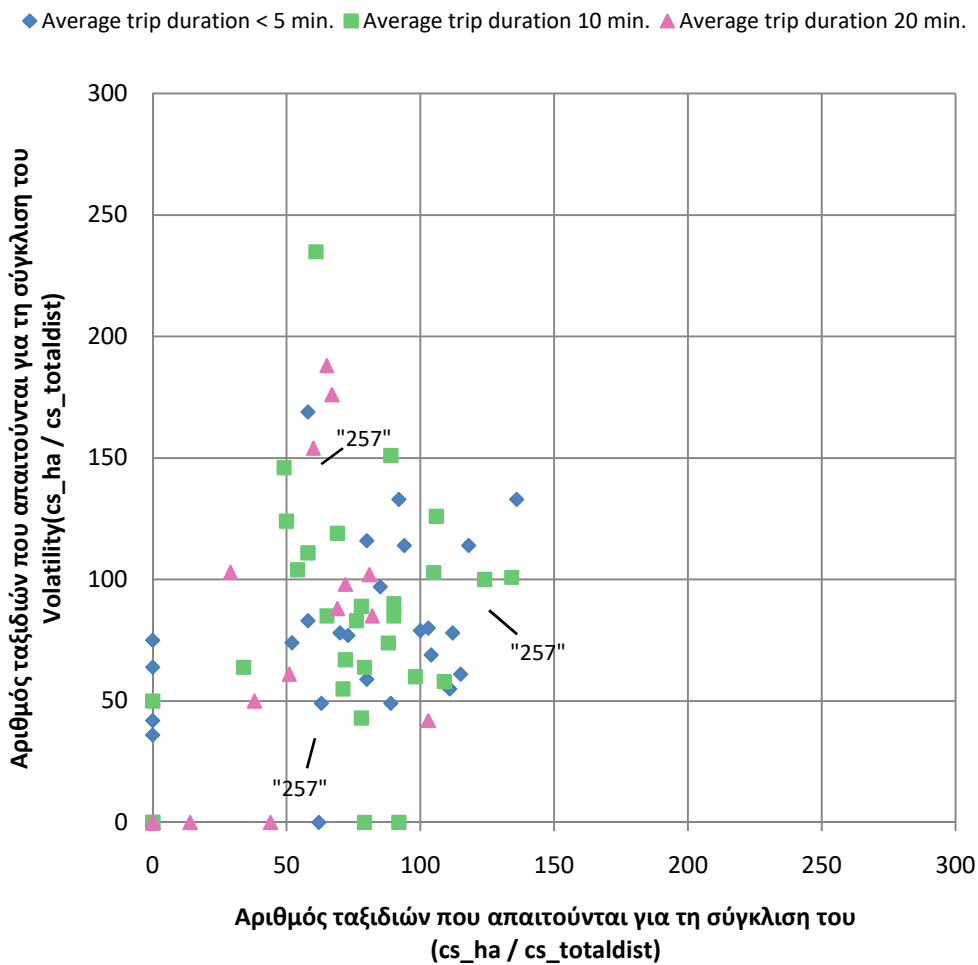
Στο Παράρτημα παρατίθενται οι συγκεντρωτικοί πίνακες **“CONVERGENCE”** όπως προέκυψαν για τους παραπάνω 29 οδηγούς, με τον πίνακα «DUR5» να αφορά τα ταξίδια όλων των οδηγών τα οποία όμως έχουν συνολική καθαρή διάρκεια οδήγησης μικρότερη των 5 λεπτών, τον «DUR10» να αφορά τα ταξίδια των οποίων η διάρκεια ήταν μεταξύ 5 και 15 λεπτών (μέση τιμή διάρκειας υποβάσης: 10 λεπτά) και τον «DUR20» ο οποίος αφορά αντίστοιχα τα ταξίδια των οποίων η διάρκεια ήταν μεταξύ 15 και 25 λεπτών (μέση τιμή διάρκειας υποβάσης: 20 λεπτά). Τα κενά κελιά σημαίνουν ότι δεν υπήρξε σύγκλιση για το εκάστοτε χαρακτηριστικό του οδηγού UserID(i) για τα ταξίδια του με διάρκεια DUR(j).

Από τα αποτελέσματα της ανάλυσης γίνεται εμφανές ότι **δεν υπάρχει ένα διακριτό χρονικό σημείο, κοινό για όλους τους οδηγούς ή/και όλα τα χαρακτηριστικά, από το οποίο και μετά η οδηγική συμπεριφορά να σταθεροποιείται**. Κάτι τέτοιο ήταν αναμενόμενο, καθώς οι οδηγοί διαφέρουν μεταξύ τους ως προς την επιθετικότητα, και συνεπώς, για τον καθορισμό του χρονικού διαστήματος παρατήρησης ενός οδηγού, θα πρέπει να έχει προηγηθεί μια ανάλυση του προφίλ επιθετικότητάς του.

Σε αυτή τη βάση αξίζει να αναφερθεί ότι, όπως προέκυψε από την ανάλυση, στους πλέον επιθετικούς οδηγούς τα χαρακτηριστικά τείνουν να συγκλίνουν με γρηγορότερο ρυθμό από ό,τι στους πιο προσεκτικούς οδηγούς, επιβεβαιώνοντας και τα αποτελέσματα της βιβλιογραφίας (Tselentis et al., (2017), Mantouka et al., (2017)).

Το χαρακτηριστικό που συγκλίνει πιο αργά για κάθε οδηγό αποτελεί και το κρίσιμο χαρακτηριστικό του οδηγού, με βάση το οποίο καθορίζεται ο ελάχιστος αριθμός ταξιδιών που απαιτείται να συλλεχθούν ώστε να είναι

εφικτή η περιγραφή της οδηγικής συμπεριφοράς του. Πολλές φορές, το μέγεθος αυτό μπορεί να αντιστοιχεί στο μέγεθος της μεταβλητότητας κάποιου χαρακτηριστικού, καθώς για τον ίδιο οδηγό, ο ρυθμός σύγκλισης ενός χαρακτηριστικού διαφέρει γενικά από τον ρυθμό σύγκλισης της μεταβλητότητας του ίδιου χαρακτηριστικού. Αυτό γίνεται πιο κατανοητό εξετάζοντας το παρακάτω διάγραμμα που αφορά τις απότομες επιταχύνσεις ανά χιλιόμετρο. Τα αντίστοιχα διαγράμματα και για τις υπόλοιπες τρεις μεταβλητές που εξετάστηκαν στην ανάλυση παρατίθενται στο Παράρτημα.



Διάγραμμα 5.1: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του μεγέθους απότομων επιταχύνσεων ανά χιλιόμετρο και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες

Στο διάγραμμα αυτό, όσο ένας οδηγός βρίσκεται πάνω στη διαγώνιο, σημαίνει ότι ο ρυθμός σύγκλισης του αριθμού απότομων επιταχύνσεων ανά

χιλιόμετρο(άξονας x) συμπίπτει (ή είναι αρκετά κοντά) με τον ρυθμό σύγκλισης της μεταβλητότητας του ίδιου μεγέθους (άξονας y). Αν ένας οδηγός βρίσκεται κάτω από την διαγώνιο, σημαίνει ότι ο ελάχιστος αριθμός ταξιδιών που πρέπει να παρατηρηθεί εξαρτάται από τον αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο(που συγκλίνει πιο αργά), ενώ αν βρίσκεται πάνω από την διαγώνιο, εξαρτάται από τη μεταβλητότητα του ίδιου μεγέθους (που συγκλίνει πιο αργά). Αντίστοιχα συμπεράσματα προκύπτουν και από την εξέταση των αντίστοιχων διαγραμμάτων (Παράρτημα) για τα υπόλοιπα μεγέθη που μελετώνται, δηλαδή των απότομων επιβραδύνσεων, του ποσοστού του χρόνου χρήσης του κινητού τηλεφώνου και του ποσοστού του χρόνου υπέρβασης του ορίου ταχύτητας κατά τη διάρκεια της οδήγησης.

Γενικά, αν κάποιος οδηγός χρειάζεται να παρατηρηθεί για περισσότερα από 120 ταξίδια μέχρι να συγκλίνει η μεταβλητότητα κάποιου χαρακτηριστικού, τότε ο ρυθμός σύγκλισης της μεταβλητότητάς του ως προς το εκάστοτε χαρακτηριστικό θεωρείται αργός. Το αντίθετο σημαίνει αν η μεταβλητότητα συγκλίνει πριν τα 60 ταξίδια παρατήρησης, οπότε και ο ρυθμός σύγκλισης θεωρείται γρήγορος. Τα όρια αυτά για τη μεταβλητότητα (volatility) των οδηγών προέκυψαν από την ανάλυση.

Η αντίστοιχη ανάλυση για τα τέσσερα χαρακτηριστικά της οδηγικής συμπεριφοράς που μελετώνται στην παρούσα εργασία, έδωσε τα εξής όρια για την γρήγορη ή αργή σύγκλισή τους, με βάση τον ελάχιστο αριθμό απαιτούμενων ταξιδιών παρακολούθησης:

Πίνακας 5.1: Διερεύνηση αργής ή γρήγορης σύγκλισης των χαρακτηριστικών οδηγικής συμπεριφοράς που μελετώνται με βάση τον ελάχιστο απαιτούμενο αριθμό ταξιδιών

	ΕΛΑΧΙΣΤΟΣ ΑΠΑΙΤΟΥΜΕΝΟΣ ΑΡΙΘΜΟΣ ΤΑΞΙΔΙΩΝ		
	Γρήγορη Σύγκλιση		Αργή Σύγκλιση
Απότομες Επιταχύνσεις ανά km	< 50	50 - 120	> 120
Απότομες Επιβραδύνσεις ανά km	< 60	60 - 140	> 140
Ποσοστό (%) χρόνου χρήσης κινητού τηλεφώνου	< 50	50 - 120	> 120
Ποσοστό (%) χρόνου υπέρβασης ορίου ταχύτητας	< 50	50 - 120	> 120
Μεταβλητότητα (Γενικά)	< 60	60 - 120	> 120

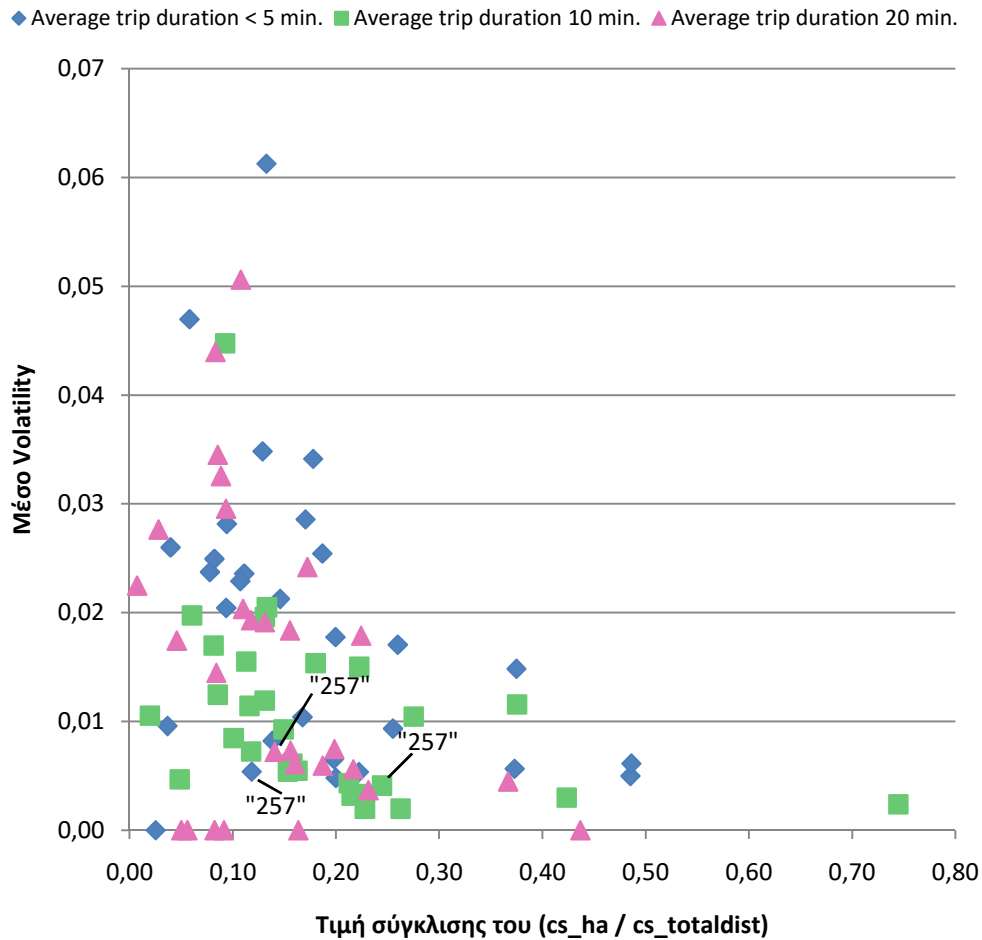
Επιπλέον, εκτός από την επιθετικότητα, ο αριθμός των ταξιδιών που απαιτείται να παρατηρηθεί ένας οδηγός διαφέρει και ως προς τη μέση χρονική διάρκεια των ταξιδιών που μελετώνται. Ακόμα και για τον ίδιο οδηγό, ο ρυθμός σύγκλισης των ίδιων χαρακτηριστικών παρουσιάζει σημαντικές διαφορές, αναλόγως με το αν μελετώνται ταξίδια μέσης χρονικής διάρκειας π.χ. 10 λεπτών ή 20 λεπτών. Στο Διάγραμμα 5.1 έχει επισημανθεί ο οδηγός με ονομασία χρήστη “257” για τις τρεις διαφορετικές διάρκειες ταξιδιών που μελετήθηκαν στην παρούσα εργασία. Γίνεται έτσι εμφανές ότι ο ίδιος οδηγός και για το ίδιο υπό μελέτη χαρακτηριστικό, αλλάζει τη σχετική θέση του στο διάγραμμα.

Από τις μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδηγικής συμπεριφοράς και η μεταβλητότητάς τους, προσδιορίζεται η επιθετικότητα και η μεταβλητότητα των οδηγών. Ένας οδηγός μπορεί να είναι προσεκτικός ως προς το χαρακτηριστικό που μελετάται (δηλαδή η μέση τιμή στην οποία συγκλίνει το μέγεθος αυτό να είναι μικρή), αλλά ταυτόχρονα να παρουσιάζει σημαντικές μεταβολές/διακυμάνσεις στη συμπεριφορά του ως προς το

εκάστοτε χαρακτηριστικό από ταξίδι σε ταξίδι (δηλαδή να έχει μεγάλο δείκτη μεταβλητότητας), ενώ μπορεί να ισχύει και το ανάποδο. Αυτό γίνεται πιο κατανοητό εξετάζοντας το Διάγραμμα 5.2 που αφορά τις απότομες επιταχύνσεις ανά χιλιόμετρο, σε συνδυασμό και με τον Πίνακα 5.2. Τα αντίστοιχα διαγράμματα και για τις υπόλοιπες τρεις μεταβλητές που εξετάστηκαν στην ανάλυση παρατίθενται στο Παράρτημα.

Πίνακας 5.2: Όρια επιθετικότητας και μεταβλητότητας οδηγών όπως προέκυψαν από την ανάλυση

ΜΕΣΕΣ ΤΙΜΕΣ ΣΥΓΚΛΙΣΗΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΚΑΙ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ			
	Προσεκτικός		Επιθετικός
Απότομες Επιταχύνσεις ανά km	<0,11	0,11–0,23	>0,23
Απότομες Επιβραδύνσεις ανά km	<0,01	0,01–0,12	>0,12
Ποσοστό (%) χρόνου χρήσης κινητού τηλεφώνου	<0,04	0,04 –0,16	>0,16
Ποσοστό (%) χρόνου υπέρβασης ορίου ταχύτητας	<0,02	0,02–0,14	>0,14
	Σταθερός		Μη Σταθερός
Μεταβλητότητα (Γενικά)	< 0,005	0,005–0,05	>0,05



Διάγραμμα 5.2: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς τον αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο

Στο Διάγραμμα 5.2 εξετάζεται η μέση τιμή της μεταβλητότητας (άξονας y) σε σχέση με τη μέση τιμή του αριθμού απότομων επιταχύνσεων ανά χιλιόμετρο (άξονας x) στις οποίες συγκλίνει κάθε οδηγός, για τις τρεις διαφορετικές διάρκειες ταξιδιών όπου πραγματοποιήθηκε η ανάλυση. Αξιοποιώντας και τα δεδομένα από τον Πίνακα 5.2 προκύπτει ότι, οι οδηγοί που βρίσκονται στην περιοχή που ορίζεται από Μέση Μεταβλητότητα του αριθμού απότομων επιταχύνσεων ανά χιλιόμετρο μικρότερη από 0,02 και τιμή σύγκλισης του αριθμού απότομων επιταχύνσεων ανά χιλιόμετρο μικρότερη από 0,17 παρουσιάζουν γενικά σταθερή συμπεριφορά (με λίγες διακυμάνσεις) από ταξίδι σε ταξίδι και ταυτόχρονα ο αριθμός των απότομων επιταχύνσεων ανά χιλιόμετρο είναι μικρός. Πρόκειται δηλαδή για την περιοχή όπου ανήκουν

οι σταθεροί και προσεκτικοί οδηγοί ως προς το χαρακτηριστικό αυτό. Αντίστοιχα, οι οδηγοί που βρίσκονται στην περιοχή που ορίζεται από μέση μεταβλητότητα μεγαλύτερη από 0,05 και αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο μικρότερο από 0,17 χαρακτηρίζονται ως μη σταθεροί, προσεκτικοί οδηγοί (ως προς το χαρακτηριστικό αυτό), ενώ εκείνοι που βρίσκονται στην περιοχή με μέση μεταβλητότητα μεγαλύτερη από 0,05 και αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο μεγαλύτερο από 0,23 χαρακτηρίζονται ως μη σταθεροί, επιθετικοί οδηγοί (ως προς το χαρακτηριστικό αυτό).

Κατά αντιστοιχία με το Διάγραμμα 5.1, η επιθετικότητα και η μεταβλητότητα ενός οδηγού φαίνεται να εξαρτάται σε ένα βαθμό και σε αυτή την περίπτωση από τη μέση χρονική διάρκεια των ταξιδιών που μελετώνται. Ακόμα και για τον ίδιο οδηγό, υπάρχουν διαφορές, αναλόγως με το αν μελετώνται ταξίδια μέσης χρονικής διάρκειας π.χ. 10 λεπτών ή 20 λεπτών, σε αντίθεση όμως με το Διάγραμμα 5.1, οι διαφορές αυτές δεν είναι συνήθως σημαντικές, δηλαδή οι οδηγοί φαίνεται να διατηρούν την ίδια κατηγορία συμπεριφοράς και μεταβλητότητας συμπεριφοράς ως προς κάποιο χαρακτηριστικό ανεξάρτητα από την μέση χρονική διάρκεια των ταξιδιών που μελετώνται. Στο Διάγραμμα 5.2 έχει επισημανθεί ο οδηγός με ονομασία χρήστη “257” για τις τρεις διαφορετικές διάρκειες ταξιδιών που μελετήθηκαν στην παρούσα εργασία. Γίνεται έτσι εμφανές κατά πόσο ο ίδιος οδηγός και για το ίδιο υπό μελέτη χαρακτηριστικό, αλλάζει τη σχετική θέση του στο διάγραμμα.

Αντίστοιχα συμπεράσματα προκύπτουν και από την εξέταση των αντίστοιχων διαγραμμάτων (Παράρτημα) για τα υπόλοιπα μεγέθη που μελετώνται, δηλαδή των απότομων επιβραδύνσεων, του ποσοστού του χρόνου χρήσης του κινητού τηλεφώνου και του ποσοστού του χρόνου υπέρβασης του ορίου ταχύτητας κατά τη διάρκεια της οδήγησης.

Στο Παράρτημα επίσης παρατίθενται ενδεικτικά διαγράμματα τυχαίων οδηγών για σύγκριση, για τα διαφορετικά χαρακτηριστικά που μελετώνται στην παρούσα εργασία και για διαφορετικές μέσες χρονικές διάρκειες ταξιδιών, στα οποία παρουσιάζεται η μεταβολή του εκάστοτε χαρακτηριστικού και της μεταβλητότητάς του, καθώς επίσης και τα χρονικά σημεία σύγκλισης.

Διερευνώντας το κρίσιμο χαρακτηριστικό για κάθε οδηγό (δηλαδή εκείνο που συγκλίνει πιο αργά), από τους 29 που τελικά χρησιμοποιήθηκαν στην ανάλυση προέκυψε ο παρακάτω συγκεντρωτικός πίνακας:

Πίνακας 5.3: Συγκεντρωτικός Πίνακας ποσοστών των οδηγών για το κρίσιμο χαρακτηριστικό τους, για κάθε διάρκεια

ΚΡΙΣΙΜΟ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ								
Απότομη Επιτάχυνση ανά km			Απότομη Επιβράδυνση ανά km		Ποσοστό (%) χρήσης κινητού τηλεφώνου		Ποσοστό (%) χρόνου υπέρβασης ορίου ταχύτητας	
Μέση								
χρονική διάρκεια ταξιδιών	Μέγεθος CumSum	Μέγεθος Volatility	Μέγεθος CumSum	Μέγεθος Volatility	Μέγεθος CumSum	Μέγεθος Volatility	Μέγεθος CumSum	Μέγεθος Volatility
< 5 min.	23.81%	48.00%	42.10%	40.00%	35.00%	16.67%	22.22%	12.50%
10 min	23.08%	32.00%	24.00%	44.40%	42.30%	19.23%	19.23%	17.39%
20 min	15.38%	54.55%	9.09%	20.00%	45.45%	23.08%	33.33%	28.57%

Από τον Πίνακα 5.3 γίνεται εμφανές ότι για την πλειοψηφία των οδηγών, κρίσιμο μέγεθος αποτελεί η μεταβλητότητα του αριθμού των απότομων επιταχύνσεων ανά km, καθώς επίσης και το ποσοστό χρόνου χρήσης κινητού τηλεφώνου κατά τη διάρκεια της οδήγησης. Ακολουθούν η απότομη επιβράδυνση ανά km και η μεταβλητότητά της, ενώ για λίγους οδηγούς φαίνεται να είναι κρίσιμο το ποσοστό χρόνου υπέρβασης ορίου ταχύτητας και η μεταβλητότητά του.

Από την παραπάνω ανάλυση εξάγονται χρήσιμα συμπεράσματα σχετικά με το αρχικό ερώτημα και τον στόχο της Διπλωματικής εργασίας. Τα συμπεράσματα αυτά παρουσιάζονται πιο αναλυτικά στο Κεφάλαιο 6.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Σύνοψη Μεθοδολογίας και Αποτελεσμάτων

Αντικείμενο της παρούσας Διπλωματικής Εργασίας αποτέλεσε η διερεύνηση της απαιτούμενης χρονικής διάρκειας παρατήρησης των οδηγών για τη συλλογή επαρκών δεδομένων οδήγησης, με σκοπό την εκτίμηση της οδηγικής τους συμπεριφοράς, αξιοποιώντας λεπτομερή δεδομένα από αισθητήρες έξυπνων κινητών τηλεφώνων. Στόχος, επομένως, ήταν η εύρεση και ταυτοποίηση ενός διακριτού χρονικού σημείου ή αριθμού ταξιδιών παρακολούθησης, από το οποίο και έπειτα η προσθήκη επιπλέον δεδομένων δεν θα προσέφερε σημαντική αύξηση της υπάρχουσας γνώσης για την συμπεριφορά του οδηγού. Τα αποτελέσματα έδειξαν ότι δεν υπάρχει κάποιο χρονικό σημείο, κοινό για όλους τους οδηγούς ή για όλα τα χαρακτηριστικά που εξετάζονται, από το οποίο και μετά η οδηγική συμπεριφορά συγκλίνει σε μια σταθερή τιμή, αλλά απαιτείται περαιτέρω διερεύνηση του προφίλ επιθετικότητας των οδηγών και του περιβάλλοντος οδήγησης.

Στο παρελθόν, το ερευνητικό ενδιαφέρον για την μελέτη της οδηγικής συμπεριφοράς υπήρξε αρκετά έντονο. Από τη βιβλιογραφική ανασκόπηση ερευνών συναφών με το αντικείμενο της παρούσας Διπλωματικής Εργασίας, σε χώρες της Ευρωπαϊκής Ένωσης, αλλά και παγκοσμίως, προέκυψε ότι μέχρι τώρα τα δεδομένα που έχουν χρησιμοποιηθεί προέρχονταν κυρίως από ερωτηματολόγια αυτοαξιολόγησης και συσκευές καταγραφής δεδομένων εγκατεστημένες εντός του οχήματος. Η εισαγωγή νέων τεχνολογιών, όπως είναι τα έξυπνα κινητά τηλέφωνα (smartphones), στο πεδίο της έρευνας των μεταφορών κατέστησε δυνατή τη συλλογή, την αποθήκευση και τη μετάδοση δεδομένων φυσικής οδήγησης με υψηλή ανάλυση. Τα δεδομένα που χρησιμοποιήθηκαν στη διπλωματική εργασία αφορούσαν σε 21.610 μετακινήσεις από 68 ξεχωριστούς οδηγούς, οι οποίοι είχαν εγκαταστήσει στο κινητό τους τηλέφωνο την εφαρμογή της OSeven Telematics.

Κατά την επεξεργασία χρησιμοποιήθηκαν διάφορα μαθηματικά και στατιστικά εργαλεία, προκειμένου να διερευνηθεί η σύγκλιση των δεδομένων χρονοσειρών και ο προσδιορισμός του χρονικού σημείου της σύγκλισης, όπως επίσης και να μελετηθεί η οδηγική συμπεριφορά. Η μεθοδολογία περιλάμβανε αρχικά τη δημιουργία μεταβλητών σωρευτικών αθροισμάτων που αφορούσαν στα γεγονότα απότομων επιταχύνσεων / επιβραδύνσεων ανά χιλιόμετρο, στο ποσοστό του χρόνου χρήσης κινητού τηλεφώνου και στο ποσοστό του χρόνου υπέρβασης του ορίου ταχύτητας, τον έλεγχο της μεταβλητότητας της συμπεριφοράς του οδηγού ως προς τα προαναφερόμενα μεγέθη, και τέλος, τη χρήση κινητών μέσων όρων των σωρευτικών αθροισμάτων και των μεταβλητοτήτων για τον προσδιορισμό της σύγκλισης και του χρονικού σημείου που συμβαίνει.

Από την παραπάνω επεξεργασία προέκυψαν τα αποτελέσματα των Συγκεντρωτικών Πινάκων που παρατίθενται στο Παράρτημα. Πιο συγκεκριμένα, από την ανάλυση των δεδομένων προέκυψε ότι ο προσδιορισμός του αριθμού των ταξιδιών που πρέπει να παρατηρηθεί ένας οδηγός προκειμένου να συγκλίνει η συμπεριφορά του ως προς κάποιο χαρακτηριστικό της οδηγικής συμπεριφοράς, εξαρτάται από την επιθετικότητα και τη σταθερότητα συμπεριφοράς του οδηγού, το περιβάλλον οδήγησης και τη μέση χρονική διάρκεια των ταξιδιών που μελετώνται. Πιο αναλυτικά, τα συμπεράσματα των αποτελεσμάτων της Διπλωματικής Εργασίας παρουσιάζονται στο επόμενο εδάφιο.

6.2 Βασικά Συμπεράσματα

Από τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας προέκυψαν αποτελέσματα άμεσα συνδεδεμένα με το αρχικό ερώτημα και στόχο της εργασίας. Στο εδάφιο αυτό επιχειρείται να δοθεί μια απάντηση στα συνολικά ερωτήματα της έρευνας με σύνθεση των αποτελεσμάτων των προηγούμενων κεφαλαίων. Τα γενικά συμπεράσματα που προέκυψαν συνοψίζονται ως εξής:

Στις έρευνες που έχουν διεξαχθεί έως τώρα παγκοσμίως αναφορικά με τη διερεύνηση της συμπεριφοράς των οδηγών, η συλλογή των δεδομένων γινόταν είτε με τη χρήση ερωτηματολογίων και πειραμάτων σε προσομοιωτές οδήγησης, είτε μέσω άλλων πολύπλοκων συνδυασμών συσκευών καταγραφής που συνδέονταν με το όχημα. Στην παρούσα Διπλωματική Εργασία εξετάστηκε η παρακολούθηση της οδηγικής συμπεριφοράς με δεδομένα που συλλέγονται από αισθητήρες έξυπνων κινητών τηλεφώνων, και διαπιστώθηκε ότι με τις σύγχρονες τεχνολογίες καταγραφής, επεξεργασίας και αποθήκευσης δεδομένων μέσω Internet, η χρήση έξυπνων κινητών τηλεφώνων μπορεί να συμβάλλει καθοριστικά στην έρευνα στον τομέα των μεταφορών. Επιπλέον της διερεύνησης της οδηγικής συμπεριφοράς, η χρήση έξυπνων κινητών τηλεφώνων μπορεί να προσφέρει επιπλέον γνώση και σχετικά με την απόσπαση της προσοχής των οδηγών.

Από την προκαταρκτική ανάλυση των δεδομένων, προέκυψε ότι ο χρόνος που απαιτείται να παρατηρηθεί ένας οδηγός προκειμένου να σταθεροποιηθεί η συμπεριφορά του ως προς κάποιο χαρακτηριστικό, εξαρτάται σε σημαντικό βαθμό από την επιθετικότητα της συμπεριφοράς του. Συγκεκριμένα, οι πλέον επιθετικοί οδηγοί τείνουν να συγκλίνουν στη μέση τιμή της συμπεριφοράς τους πιο γρήγορα σε σύγκριση με τους πιο προσεκτικούς οδηγούς.

Επιθετικοί θεωρούνται οι οδηγοί με υψηλό αριθμό απότομων συμβάντων και υπέρβασης του ορίου ταχύτητας. Από την ανάλυση προέκυψε ότι οι ίδιοι οδηγοί που παρουσιάζουν υψηλές μέσες τιμές σύγκλισης απότομων επιταχύνσεων ανά χιλιόμετρο παρουσιάζουν και υψηλές μέσες τιμές απότομων επιβραδύνσεων ανά χιλιόμετρο, ενώ αντίστοιχα, οι ίδιοι που παρουσιάζουν χαμηλές μέσες τιμές απότομων επιταχύνσεων ανά χιλιόμετρο παρουσιάζουν και χαμηλές μέσες τιμές απότομων επιβραδύνσεων ανά χιλιόμετρο.

Εκτός από την επιθετικότητα όμως, άλλο ένα χαρακτηριστικό των οδηγών το οποίο επηρεάζει τον χρόνο σύγκλισης, είναι και η σταθερότητα ή μεταβλητότητα της συμπεριφοράς. Κατά τη μελέτη της οδηγικής συμπεριφοράς, η γνώση της μεταβλητότητας της συμπεριφοράς των οδηγών

είναι εξαιρετικής σημασίας, καθώς παρέχει σημαντικές πληροφορίες για την εμπειρία και τον τρόπο οδήγησης από ταξίδι σε ταξίδι. Στην διερεύνηση του κρίσιμου χαρακτηριστικού παρατήρησης κάθε οδηγού, δηλαδή εκείνου που συγκλίνει με τον πιο αργό ρυθμό, στην πλειοψηφία των περιπτώσεων η μεταβλητότητα της συμπεριφοράς αποτελούσε το κρίσιμο μέγεθος παρατήρησης, αγγίζοντας το 54,55% για το μέγεθος των ha/km στα ταξίδια μέσης διάρκειας 20 λεπτών .

Η ταυτοποίηση του χρονικού σημείου σύγκλισης της συμπεριφοράς ενός οδηγού ως προς κάποιο χαρακτηριστικό, είναι εξαιρετικά δύσκολη όταν τα ταξίδια που μελετώνται δεν έχουν την ίδια ή παρόμοια διάρκεια. Καθώς τα δεδομένα που παρέχονται από τους αισθητήρες έξυπνων κινητών τηλεφώνων αφορούν σε χαρακτηριστικά οδηγικής συμπεριφοράς σε επίπεδο ταξιδιού, η ανάλυσή τους και ο έλεγχος της μεταβολής τους σαν χρονοσειρά δεδομένων δεν μπορεί να δώσει σωστά αποτελέσματα, εφόσον η μέση διάρκεια μεταξύ διαδοχικών ταξιδιών μπορεί να διαφέρει κατά πολύ. Απαιτείται λοιπόν η ομαδοποίηση των ταξιδιών ώστε να παρουσιάζουν παρόμοια διάρκεια με ταυτόχρονη ταξινόμηση σε χρονολογική σειρά.

Η χρονική διάρκεια των ταξιδιών που αναλύονται επηρεάζει επίσης το χρονικό σημείο σύγκλισης της συμπεριφοράς ενός οδηγού. Συγκεκριμένα προέκυψε ότι, ο ίδιος οδηγός μπορεί να παρουσιάζει πολύ διαφορετικό ρυθμό συγκλίνουσας συμπεριφοράς ως προς κάποιο συγκεκριμένο χαρακτηριστικό, όταν εξετάζονται ταξίδια του με διαφορετική μέση διάρκεια.

Η χρονική διάρκεια των ταξιδιών που αναλύονται δεν επηρεάζει σημαντικά τη μέση τιμή σύγκλισης της συμπεριφοράς ενός οδηγού ως προς κάποιο χαρακτηριστικό. Συγκεκριμένα προέκυψε ότι ο ίδιος οδηγός, όταν εξετάζεται η συμπεριφορά του ως προς ένα συγκεκριμένο χαρακτηριστικό, δεν παρουσιάζει αξιόλογες μεταβολές στην τιμή στην οποία συγκλίνει αυτό το χαρακτηριστικό, όταν αλλάζει η μέση διάρκεια των ταξιδιών που μελετώνται. Οι οδηγοί δηλαδή παρουσιάζουν παρόμοια συμπεριφορά, είτε διανύουν σύντομες διαδρομές, είτε λίγο μεγαλύτερες. Αυτό ίσως οφείλεται στο γεγονός ότι οι οδηγοί που συμμετείχαν στο πείραμα, πραγματοποιούσαν επί το πλείστον αστικές

διαδρομές με τις οποίες ήταν εξοικειωμένοι, και συνεπώς το οδικό περιβάλλον και οι συνθήκες οδήγησης τους ήταν οικεία.

6.3 Διατύπωση Προτάσεων

Σύμφωνα με τα αποτελέσματα και τα συνολικά συμπεράσματα που εξήχθησαν κατά την εκπόνηση της Εργασίας αυτής, επιχειρείται η παράθεση μιας σειράς προτάσεων, οι οποίες ενδεχομένως να συμβάλουν στη βελτίωση της κυκλοφορίας και της οδικής ασφάλειας.

Η χρήση των υπηρεσιών καταγραφής και αυτόματης ενημέρωσης της οδηγικής συμπεριφοράς μέσω της χρήσης έξυπνων κινητών τηλεφώνων, όπως στην παρούσα Διπλωματική Εργασία, θα επιτρέψει ενδεχομένως, όλο και περισσότεροι οδηγοί να ενημερώνονται και να βελτιώνουν μεμονωμένα τη συμπεριφορά τους, αλλά και συνολικά την οδική ασφάλεια.

Η παρακίνηση των οδηγών να αλλάξουν τον τρόπο με τον οποίο οδηγούν και να υιοθετήσουν πιο βιώσιμες συμπεριφορές είναι μία από τις μεγαλύτερες προκλήσεις στις μεταφορές. Με τη χρήση της Παιχνιδοποίησης (Gamification), κατά την οποία χρησιμοποιούνται διάφορες προσεγγίσεις και τεχνικές για την ενεργοποίηση αλλαγής συμπεριφοράς και την αύξηση της συμμετοχής των οδηγών, είναι δυνατό να επιτευχθεί ο στόχος αυτός. Στις περιπτώσεις που εντοπίζεται μη ασφαλής οδηγική συμπεριφορά, τα έξυπνα κινητά τηλέφωνα μπορούν να χρησιμοποιηθούν για την υποβολή συστάσεων προς τους οδηγούς προκειμένου να οδηγηθούν σε ασφαλέστερες συμπεριφορές οδήγησης.

Θα μπορούσαν επίσης να χρησιμοποιηθούν υποδείξεις που αφορούν στην κατανάλωση καυσίμων ανάλογα με την οδηγική συμπεριφορά και η παρακίνηση μέσω συστάσεων προς τους οδηγούς, να μειώσουν τα απότομα συμβάντα και την ταχύτητά του, ώστε να εξοικονομούν καύσιμα αλλά και να μειώνουν τις εκπομπές ρύπων στην ατμόσφαιρα.

Ακόμα, θα πρέπει να υπάρξει συντονισμένη αντιμετώπιση του προβλήματος και από την πολιτεία αλλά και πολύ περισσότερο από τους εκπαιδευτικούς φορείς, έτσι ώστε να δημιουργηθεί το υπόβαθρο για τη δημιουργία υπεύθυνων και υποδειγματικών οδηγών από μικρή ηλικία. Είναι λοιπόν απαραίτητη η ανάπτυξη εκπαιδευτικών προγραμμάτων, αλλά και η ένταξη μαθημάτων οδικής ασφάλειας στο πλαίσιο των σχολικών δραστηριοτήτων, με ειδική αναφορά στους κινδύνους που εγκυμονεί η επιθετική οδήγηση, η μη συμμόρφωση στον Κ.Ο.Κ και η απόσπαση προσοχής (π.χ. χρήση του κινητού τηλεφώνου) κατά την οδήγηση.

Τέλος, τα αποτελέσματα της παρούσας Διπλωματικής Εργασίας θα μπορούσαν να χρησιμοποιηθούν από τις ασφαλιστικές εταιρείες οχημάτων, παρακινώντας τους μελλοντικούς πελάτες τους να χρησιμοποιήσουν μια τέτοια εφαρμογή, είτε για ένα προκαθορισμένο διάστημα ώστε να καταγραφεί η οδηγική συμπεριφορά τους και έπειτα να προσαρμοστούν σε αυτή τα ασφάλιστρα (παρέχοντας μια έκπτωση στα ασφάλιστρα), είτε για όλη τη διάρκεια του συμβολαίου προκειμένου να επιβραβεύουν τους προσεκτικούς οδηγούς με συνεχώς μειωμένα ασφάλιστρα όσο βελτιώνουν τη συμπεριφορά τους. Αυτή η τακτική ενδεχομένως θα δώσει επιπλέον οικονομικό κίνητρο σε όλους τους οδηγούς με στόχο τη βελτίωση της οδηγικής συμπεριφοράς και συνεπώς τη μείωση των ατυχημάτων.

Επιπροσθέτως, το θέμα της διπλωματικής εργασίας που παρουσιάστηκε επιδέχεται αρκετές αλλαγές, οι οποίες παρατίθενται παρακάτω σαν προτάσεις για περαιτέρω εμβάθυνση και έρευνα.

Όσον αφορά στη μεθοδολογία ανάλυσης, ενδιαφέρον θα παρουσίαζε η εφαρμογή διαφορετικών μεθόδων στατιστικής ανάλυσης, όπως η ανάλυση ομαδοποίησης (k-means clustering), η διερεύνηση της επιρροής των ακραίων τιμών (outliers) στα αποτελέσματα της ομαδοποίησης K-μέσων και η ανάλυση παραγόντων με σκοπό την ομαδοποίηση των οδηγών ανάλογα με τη συμπεριφορά τους.

Η παρατήρηση των ίδιων μεταβλητών σε μεγαλύτερο δείγμα οδηγών, θα παρουσίαζε επίσης αρκετό ενδιαφέρον. Συγκεκριμένα, όσο περισσότεροι οδηγοί συμμετέχουν στο πείραμα, τόσο πιο αξιόπιστα αποτελέσματα προκύπτουν και ίσως αυτό να έδινε τη δυνατότητα εξαγωγής περισσότερο ακριβών αποτελεσμάτων, κάνοντας πιο προφανή την ομαδοποίηση των οδηγών.

Ακόμα, ενδιαφέρον θα παρουσίαζε μία ανάλυση η οποία θα στηριζόταν σε περισσότερα δημογραφικά χαρακτηριστικά των οδηγών, όπως το φύλλο, η ηλικία, η εμπειρία αλλά και τα χαρακτηριστικά του οχήματος(κινητήριος δύναμη, ίπποι, παλαιότητα κλπ.) ώστε να μπορούν να ομαδοποιηθούν οι συμμετέχοντες και με βάση τα δημογραφικά χαρακτηριστικά και να προκύψουν πιο λεπτομερή αποτελέσματα.

Τέλος, ενδιαφέρουσες επίσης θα ήταν και έρευνες αντίστοιχες της παρούσας, οι οποίες μπορούν να πραγματοποιηθούν χρησιμοποιώντας ως παραμέτρους τις διαφορετικές συνθήκες κυκλοφορίας και διαφορετικά οδικά περιβάλλοντα (υψηλή/χαμηλή κυκλοφορία, ημέρα/νύχτα, διάφορες καιρικές συνθήκες κα.), αλλά και διάφορες ομάδες οδηγών (επαγγελματίες οδηγοί, μοτοσυκλετιστές, κλπ.).

ΒΙΒΛΙΟΓΡΑΦΙΑ

G. Albert, O. Musicant, T. Lotan, T. Toledo, E. Grimberg, "Evaluating Changes in the Driving Behavior of Young Drivers a Few Years After Licensure Using In-Vehicle Data Recorders", *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 2011*

P. Baecke, L. Bocca, "The value of vehicle telematics data in insurance risk selection processes", *Decision Support Systems 98 (2017), p. 69–79*

P.Bonsall, R. Liu, W. Young, Modelling safety-related driving behaviour—impact of parameter values", *Transportation Research Part A 39 (2005), p. 425–444*

M.H. Camerona, R. Elvikb, "Nilsson's Power Model connecting speed and road trauma: Applicability by road and alternative models for urban roads", *Accident Analysis and Prevention 42 (2010), p. 1908–1915*

L. Eboli, G. Mazzulla, G. Pungillo, "Combining speed and acceleration to define car users' safe or unsafe driving behavior", *Transportation Research Part C 68 (2016), p. 113–125*

A. B. Ellison, S. P. Greaves, M. C.J. Bliemer, "Driver behaviour profiles for road safety analysis", *Accident Analysis and Prevention 76 (2015), p. 118–132*

M. M. Haque, S. Washington, " The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model", *Transportation Research Part C 50 (2015), p. 13-27*

P. Händel, I. Skog, J. Wahlström, F. Bonawiede, R. Welch, J. Ohlsson, M. Ohlsson, "Insurance telematics: opportunities and challenges with the

smartphone solution”, *IEEE Intelligent Transportation Systems Magazine*, July 24, 2014

J. W. Joubert, Dirk de Beer, Nico de Koker, “Combining accelerometer data and contextual variables to evaluate the risk of driver behavior”, *Transportation Research Part F 41* (2016), p. 80–96

E. G. Mantouka, E. N. Barmounakis, E. I. Vlahogianni, “Mobile Sensing and Machine Learning for Identifying Driving Safety Profiles”, *Transportation Research Board 97th Annual Meeting*, 2018

O. Musicant, H. Bar-Gera, E. Schechtman, “Individual Driver’s Undesirable Driving Events – A Temporal Analysis”, *3rd International Conference on Road Safety and Simulation*, September 14-16, 2011, Indianapolis, USA

O. Musicant, H. Bar-Gera, E. Schechtman, “Electronic records of undesirable driving events”, *Transportation Research Part F 13* (2010), p. 71–79

O. Musicant, T. Lotan, T. Toledo, “Safety correlation and implications of an in-vehicle data recorder on driver behavior”, *TRB 2007 Annual Meeting CD-ROM*

P. Nitsche, P. Widhalm, S. Breuss, N. Brändle, P. Maurer, “Supporting large-scale travel surveys with smartphones – A practical approach”, *Transportation Research Part C 43* (2014), p. 212–221

A. Pakgohar, R. S. Tabrizi, M. Khalili, A. Esmaeili, “The role of human factor in incidence and severity of road crashes based on the CART and LR regression: data mining approach”, *Procedia Computer Science 3* (2011), p. 764–769

R. Paleti, O. Sahin, M. Cetin, “Modeling the impact of latent driving patterns on traffic safety using mobile sensor data”, *Accident Analysis and Prevention 107* (2017), p. 92–101

M. Papadakaki, G. Tzamalouka, C. Gnardellis, T. J. Lajunen, J. Chliaoutakis, “Driving performance while using a mobile phone: A simulation study of Greek professional drivers”, *Transportation Research Part F* 38 (2016), p. 164–170

R. Shichrur, A. Sarid, N. Z. Ratson, “Determining the sampling time frame for In-Vehicle Data Recorder measurement in assessing drivers”, *Transportation Research Part C* 42 (2014), p. 99–106

Y. Shiftan, G. Toledo, “Analysis and Modeling of Driving Behavior using In-Vehicle Data Recorder”, Transportation Research Institute, TechnionTechnion City, Haifa 2007

D. Shinar, R. Compton, “Aggressive driving: an observational study of driver, vehicle, and situational variables”, 2004

G. Singh, D. Bansal, S. Sofat, “A smartphone based technique to monitor driving behavior using DTW and crowdsensing”, *Pervasive and Mobile Computing* 40 (2017), p. 56–70

A. Theofilatos, D. Tselentis, G. Yannis, M. Konstantinopoulos, “Willingness-to-Pay for Usage-Based Motor Insurance”, *TRB 2017 Annual Meeting*

T. Toledo, O. Musicant, T. Lotan, “In-vehicle data recorders for monitoring and feedback on drivers’ behavior”, *Transportation Research Part C* 16 (2008), p. 320–331

T. Toledo, T. Lotan, “An In-Vehicle Data Recorder for Evaluation of Driving Behavior and Safety”, *TRB 2006 Annual Meeting CD-ROM*

G. Toledo, Y. Shiftan, “Can feedback from in-vehicle data recorders improve driver behavior and reduce fuel consumption?”, *Transportation Research Part A* 94 (2016), p.194–204

D. I. Tselentis, E. I. Vlahogianni, G. Yannis, “Quantifying the Need for Driving Data Collection in Driving Behaviour Assessment Using Smartphone Data”, *Proceedings of 7th Transport Research Arena TRA 2018*, Vienna, Austria, April 16-19, 2018

D. I. Tselentis, E. I. Vlahogianni, M. G. Karlaftis, “Improving short-term traffic forecasts: to combine models or not to combine?”, *IET Intelligent Transport Systems 2013*, pp. 1–9

D. I. Tselentis, G. Yannis, E. I. Vlahogianni, “Innovative motor insurance schemes: A review of current practices and emerging challenges”, *Accident Analysis and Prevention 98* (2017), p. 139–148

E.I. Vlahogianni, E.N. Barmounakis, “Driving analytics using smartphones: Algorithms, comparisons and challenges”, *Transportation Research Part C 79* (2017), p. 196–206

X. Zhu, X. Hu, Y. Chiu, “Design of Driving Behavior Pattern Measurements Using Smartphone Global Positioning System Data”, *International Journal of Transportation Science and Technology vol. 2, no. 4* (2013), p. 269–288

Φ.Αδαμίδης, “Διερεύνηση των Επιδράσεων Προτύπων Οικολογικής και Απότομης Οδήγησης στα Μεγέθη της Κυκλοφοριακής Ροής και στις Εκπομπές Ρύπων”, *Διπλωματική Εργασία, Σχολή Πολιτικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Οκτώβριος 2017*

A. Αργυροπούλου, “Προτυποποίηση της Επιρροής της Χρήσης Κινητού Τηλεφώνου στη Συμπεριφορά του Οδηγού Αξιοποιώντας Λεπτομερή Δεδομένα από Αισθητήρες Έξυπνων Κινητών Τηλεφώνων”, *Διπλωματική Εργασία, Σχολή Πολιτικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Ιούλιος 2017*

X. Γονίδη, “Ανάπτυξη Προτύπων Ταχύτητας του Οδηγού με βάση Λεπτομερή Δεδομένα Οδήγησης από Αισθητήρες Κινητών Τηλεφώνων”, *Διπλωματική*

Εργασία, Σχολή Πολιτικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Ιούλιος 2017

Α. Λ. Στεφανίδης, “Μοντέλα Ακολουθιακής Στατιστική Ανάλυσης και Εφαρμογές”, *Διπλωματική Εργασία, Σχολή Χρηματοοικονομικής και Στατιστικής, Πανεπιστήμιο Πειραιώς, Πειραιάς, Σεπτέμβριος 2015*

Δ. Τσελέντης, “Πρόβλεψη Ταχύτητας σε Ελεύθερες Λεωφόρους με Μοντέλα Χρονοσειρών με Χωρο-Χρονικές Συσχετίσεις”, *Διπλωματική Εργασία, Σχολή Πολιτικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Οκτώβριος 2012*

OSeven, 2018, www.oseven.io, Ανάκτηση 20/05/2018

WHO (World Health Organization), Global Status Report on Road Safety, 2015, www.who.int/violence_injury_prevention/road_safety_status/2015/en Ανάκτηση 12/05/2018

Ελληνική Αστυνομία,
www.astynomia.gr/index.php?option=ozo_content&perform=view&id=74456&Itemid=86&lang= , Ανάκτηση 20/05/2018

ΠΑΡΑΡΤΗΜΑ

Πίνακας 0.1: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να συλλεχθούν για να επιτευχθεί σύγκλιση (Ταξίδια μέσης διάρκειας < 5 min.)

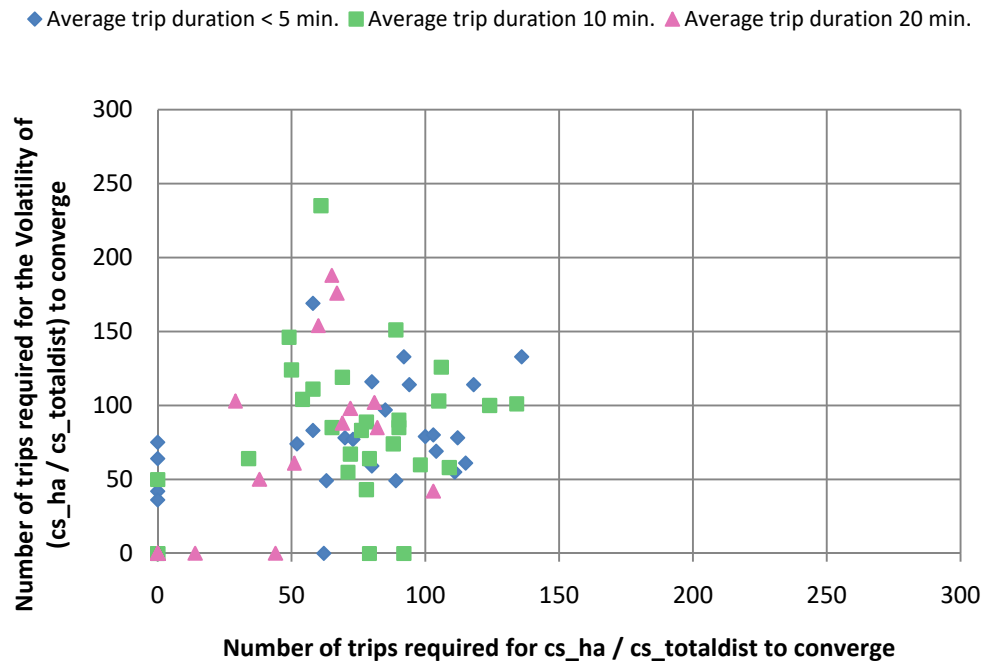
USER_ID	DUR 5							
	HARSH ACCELERATION		HARSH BRAKING		TIME MOBILE USAGE		SUM SPEEDING	
	cs_ha / totaldist	volatility	cs_hb / totaldist	volatility	cs_TMU / duration	volatility	cs_sum_speeding / duration	volatility
3		64		85		76		76
4	100	79	85	51	92	59	138	157
9	58	169	95	63	106	43	93	134
10	104	69	157	107	89	66	69	68
11	62		56	67		52		
18	136	133	93	70	59	112	76	53
19								
25	52	74	98	54	77	97	106	51
35	70	78		52	73	66	74	40
46	80	116	60	50	59	79	145	70
94	115	61	102	58	104	74	77	77
119	89	49				53	64	47
124	80	59	117	103		77	81	70
132		42				38		
140	94	114	196	271	94	72	113	55
143	92	133	146	83	95	93	94	89
146	118	114	119	73	79	77	97	97
147	85	97	84	63	103			66
149								
154	103	80			167	187	142	172
159	58	83	123	82	89	63	110	77
257	63	49	63	82	76	73	65	62
332	112	78	83		114	58	74	41
374			69		69	67		41
427	111	55	109	97	92	80		61
436	73	77	79	98	69	58	70	47
465		75		67	52	69		50
466								34
553		36						

Πίνακας 0.2: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να συλλεχθούν για να επιτευχθεί σύγκλιση (Ταξίδια μέσης διάρκειας 10 min.)

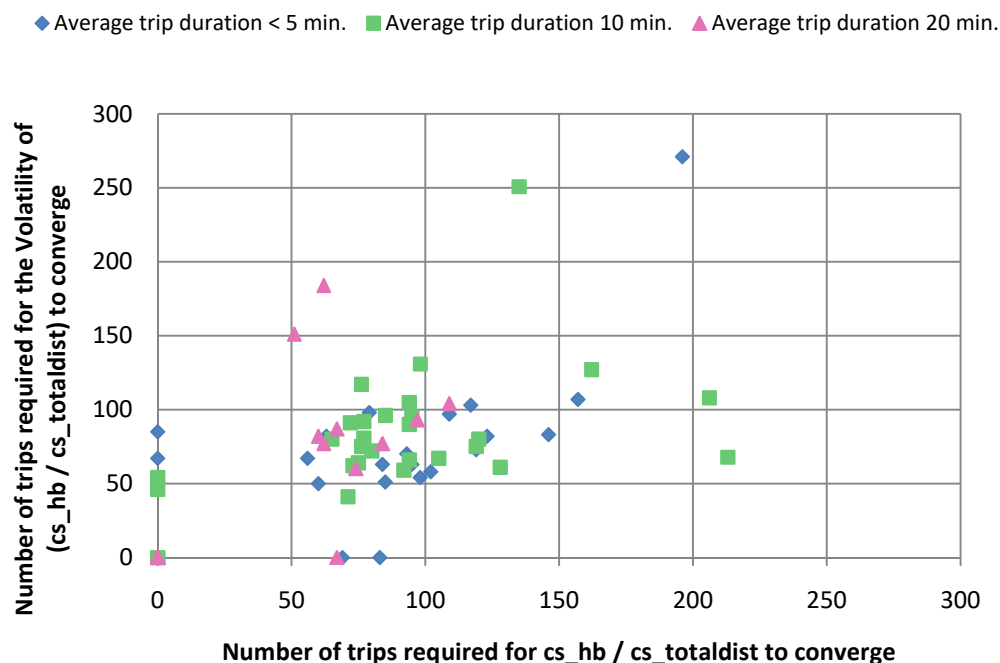
USER_ID	DUR 10							
	HARSH ACCELERATION		HARSH BRAKING		TIME MOBILE USAGE		SUM SPEEDING	
	cs_ha / totaldist	volatility	cs_hb / totaldist	volatility	cs_TMU / duration	volatility	cs_sum_speeding / duration	volatility
3	76	83	73	62	152	71	63	83
4	65	85	94	90	134	79	97	69
9	78	89	120	80	128	52	99	83
10	49	146	76	117	98	60	88	
11	92		72	91	81	63	53	
18	61	235	94	105	123	72	120	64
19		50				52		
25	54	104	135	251	74	104	91	171
35	71	55	85	96	85	105	105	77
46	58	111	80	72	70	62	123	201
94	134	101	95	95	99	73	101	71
119	98	60	71	41	99	61	54	78
124	79			54	79		76	
132	34	64	92	59	41	77	18	41
140	106	126	98	131	99	85	89	85
143	50	124	65	80	115	100	154	157
146	88	74	119	75	104	58	131	62
147	79	64	77	92	87	73	86	86
149	72	67	76	75	79	55	81	55
154	90	85	206	108	291	88	76	74
159	105	103	213	68	86	80	108	91
257	124	100	94	66	119	83	87	76
332	90	90	162	127	82	63	70	79
374	78	43	77	81	101		63	99
427	109	58	75	64	84	46	82	51
436	69	119	105	67	79	68	57	48
465					67	63		
466				46				
553	89	151	128	61		203	89	67

Πίνακας 0.3: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να συλλεχθούν για να επιτευχθεί σύγκλιση (Ταξίδια μέσης διάρκειας 20 min.)

USER_ID	DUR 20							
	HARSH ACCELERATION		HARSH BRAKING		TIME MOBILE USAGE		SUM SPEEDING	
	cs_ha / totaldist	volatility	cs_hb / totaldist	volatility	cs_TMU / duration	volatility	cs_sum_speeding / duration	volatility
3								
4	81	102	74	60	58	69	106	85
9	44		60	82	84	118	97	84
10								
11	29	103	67		73	117	36	50
18	51	61	51	151	103	52	61	51
19	69	88	84	77	97	34	56	76
25								
35								
46								46
94						41		52
119								53
124								
132								
140	82	85	97	93	143	94	80	166
143	65	188	62	77	156	53	101	83
146	67	176	67	87	86	60	74	140
147								
149								
154	103	42	102		96	116	102	50
159	72	98	62	184	72	44	126	40
257	60	154	109	104	92	92	74	108
332	38	50				38	59	
374								
427	14							
436								
465								
466								
553								

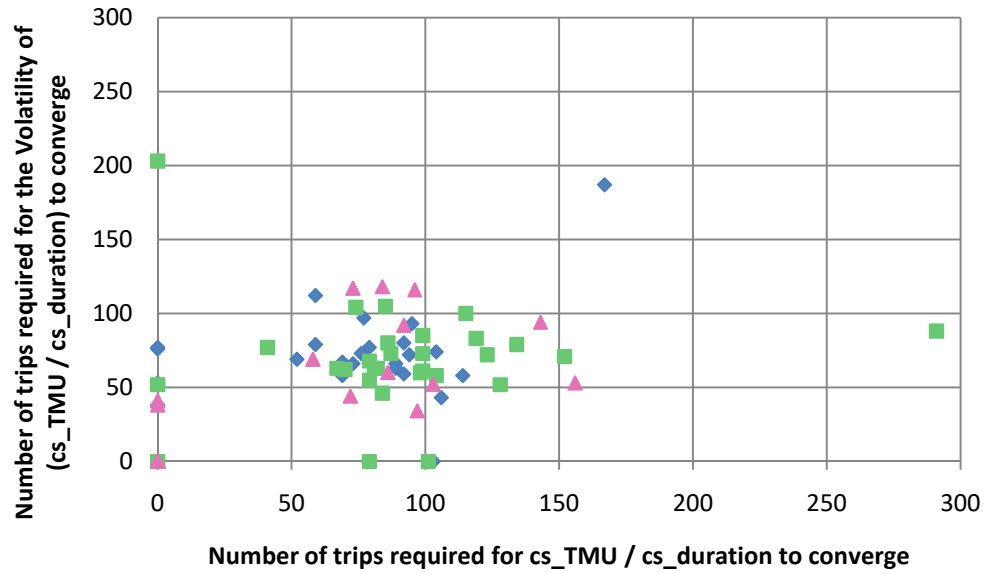


Διάγραμμα 0.1: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του μεγέθους απότομων επιταχύνσεων ανά χιλιόμετρο και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες



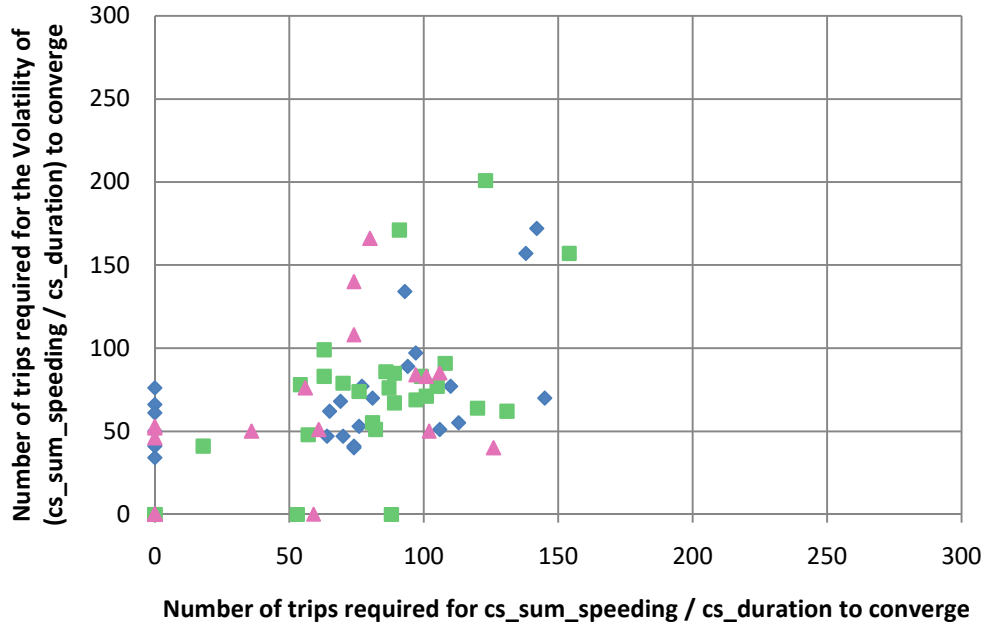
Διάγραμμα 0.2: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του μεγέθους απότομων φρεναρισμάτων ανά χιλιόμετρο και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες

◆ Average trip duration < 5 min. ■ Average trip duration 10 min. ▲ Average trip duration 20 min.



Διάγραμμα 0.3: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του ποσοστού χρήσης κινητού τηλεφώνου και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες

◆ Average trip duration < 5 min. ■ Average trip duration 10 min. ▲ Average trip duration 20 min.



Διάγραμμα 0.4: Ελάχιστος απαιτούμενος αριθμός ταξιδιών που πρέπει να καταγραφούν για κάθε οδηγό με βάση τη σύγκλιση του ποσοστού του χρόνου υπέρβασης ορίου ταχύτητας και με βάση τη μεταβλητότητα του ίδιου μεγέθους, για τις τρεις διαφορετικές διάρκειες

Πίνακας 0.4: Μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδήγικής συμπεριφοράς (Ταξίδια μέσης χρονικής διάρκειας < 5 min.)

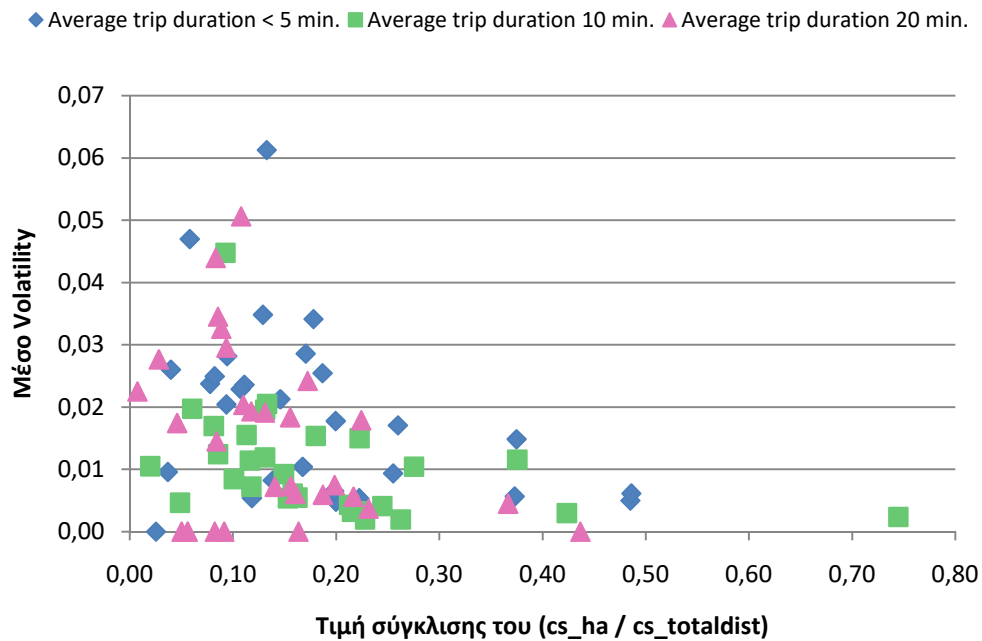
USER_ID	DUR 5							
	HARSH ACCELERATION		HARSH BRAKING		TIME MOBILE USAGE		SUM SPEEDING	
	cs_ha / totaldist	volatility	cs_hb / totaldist	volatility	cs_TMU / duration	volatility	cs_sum_speeding / duration	volatility
3	0.07783	0.02373	0.04540	0.02768	0.01217	0.06907	0.01274	0.03711
4	0.37293	0.00564	0.11804	0.00830	0.10426	0.01256	0.09852	0.00524
9	0.48510	0.00499	0.05505	0.01259	0.02940	0.01215	0.08793	0.00993
10	0.22224	0.00536	0.05983	0.00727	0.09003	0.00727	0.02210	0.00520
11	0.37473	0.01484	0.17565	0.01839	0.15438	0.04099	0.21331	0.01619
18	0.25513	0.00934	0.11072	0.01119	0.24155	0.01113	0.03441	0.01675
19	0.02540		0.02540	0.01534	0.01769	0.06303	0.00099	
25	0.48595	0.00613	0.18371	0.01085	0.33417	0.00455	0.07566	0.00726
35	0.18676	0.02544	0.04150	0.04034	0.15629	0.02559	0.04714	0.03871
46	0.11829	0.00539	0.04301	0.01060	0.13525	0.00873	0.02306	0.01163
94	0.19946	0.01775	0.10228	0.01446	0.03953	0.03275	0.04451	0.02421
119	0.11109	0.02359	0.03471	0.03099	0.13292	0.04796	0.07617	0.02879
124	0.09365	0.02044	0.03512	0.04882	0.32543	0.01528	0.02911	0.01888
132	0.13264	0.06128	0.04421	0.06216	0.03003	0.06168	0.02002	0.08054
140	0.03686	0.00958	0.00890	0.00720	0.10204	0.00651	0.01219	0.00895
143	0.19979	0.00481	0.10632	0.00923	0.09183	0.00661	0.05700	0.00504
146	0.19792	0.00653	0.08259	0.00725	0.15734	0.00729	0.02447	0.00664
147	0.17044	0.02857	0.14845	0.01824	0.10901	0.01465	0.03362	0.03072
149	0.17792	0.03414	0.09364	0.05174	0.03997	0.14539	0.18332	0.03473
154	0.03976	0.02600	0.00000		0.01708	0.02270	0.00803	0.03981
159	0.16754	0.01040	0.05046	0.01688	0.05663	0.01128	0.01656	0.01037
257	0.13850	0.00822	0.06575	0.00936	0.10020	0.01158	0.05872	0.00981
332	0.14585	0.02127	0.05014	0.03513	0.01256	0.02518	0.04297	0.02255
374	0.12893	0.03483	0.10618	0.04332	0.25223	0.01720	0.06905	0.04433
427	0.08223	0.02494	0.01028	0.02125	0.02485	0.04815	0.01094	0.04077
436	0.25977	0.01705	0.02969	0.03086	0.17843	0.01607	0.06611	0.02501
465	0.10738	0.02289	0.10022	0.02549	0.14743	0.02768	0.01019	0.04782
466	0.09429	0.02816	0.02829	0.07419	0.30504	0.01360	0.06447	0.05317
553	0.05812	0.04698	0.00000		0.00000		0.00180	

Πίνακας 0.5: Μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδηγικής συμπεριφοράς (Ταξίδια μέσης χρονικής διάρκειας 10 min.)

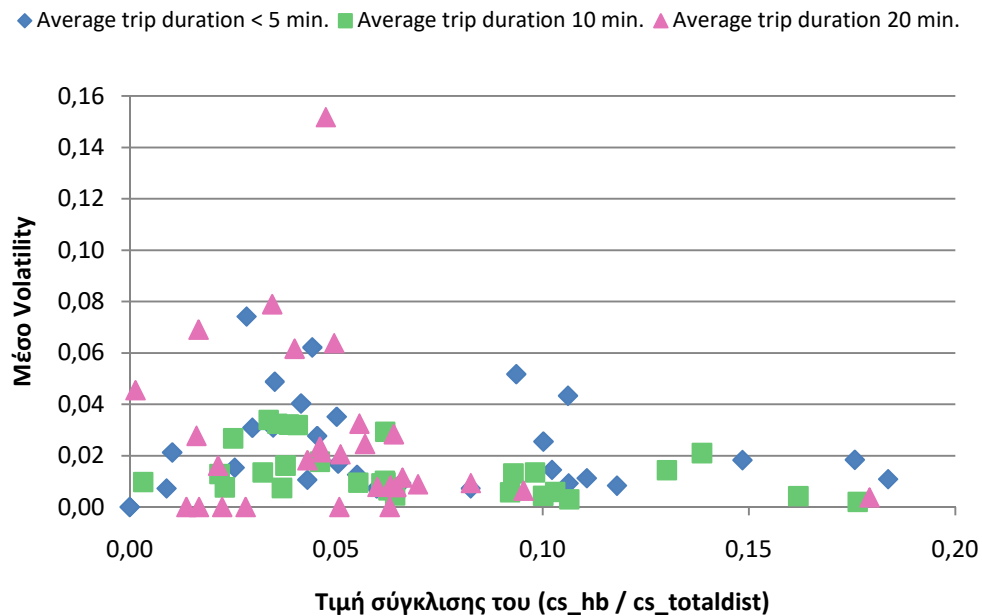
USER_ID	DUR 10							
	HARSH ACCELERATION		HARSH BRAKING		TIME MOBILE USAGE		SUM SPEEDING	
	cs_ha / totaldist	volatility	cs_hb / totaldist	volatility	cs_TMU / duration	volatility	cs_sum_speeding / duration	volatility
3	0.11770	0.00724	0.06195	0.00954	0.01436	0.00858	0.03187	0.00891
4	0.24457	0.00411	0.10306	0.00584	0.14629	0.00712	0.14131	0.00519
9	0.42350	0.00300	0.06265	0.00658	0.02910	0.01160	0.09313	0.00370
10	0.21994	0.00331	0.10014	0.00439	0.04846	0.00876	0.11087	0.00386
11	0.27529	0.01044	0.09819	0.01349	0.16107	0.02361	0.26349	0.00693
18	0.26244	0.00199	0.17631	0.00204	0.24099	0.00431	0.05729	0.00414
19	0.08547	0.01247	0.06189	0.02930	0.02764	0.04288	0.04743	0.01430
25	0.74481	0.00238	0.16194	0.00422	0.16394	0.00502	0.12557	0.00234
35	0.13084	0.01193	0.03772	0.01615	0.12159	0.01287	0.05069	0.01776
46	0.15325	0.00538	0.03688	0.00743	0.07130	0.00553	0.06362	0.00597
94	0.15773	0.00612	0.06180	0.01025	0.02294	0.00903	0.05066	0.00820
119	0.06039	0.01975	0.04067	0.03186	0.12286	0.04321	0.11063	0.04148
124	0.11297	0.01552	0.03577	0.03241	0.23704	0.01987	0.15062	0.02522
132	0.11613	0.01144	0.02507	0.02674	0.04648	0.00902	0.04263	0.02315
140	0.04849	0.00467	0.02305	0.00767	0.04628	0.00463	0.03757	0.00687
143	0.22797	0.00198	0.10648	0.00304	0.06998	0.00401	0.05597	0.00227
146	0.21241	0.00433	0.09213	0.00577	0.11905	0.00815	0.04405	0.00546
147	0.18018	0.01537	0.13013	0.01431	0.16930	0.01345	0.04634	0.01538
149	0.22258	0.01502	0.13859	0.02101	0.02727	0.02496	0.09765	0.01306
154	0.01969	0.01053	0.00328	0.00970	0.03649	0.00939	0.02057	0.02046
159	0.16249	0.00548	0.06093	0.00926	0.03709	0.00865	0.03051	0.00829
257	0.21506	0.00320	0.06427	0.00447	0.08144	0.00377	0.04596	0.00416
332	0.14917	0.00927	0.05536	0.00944	0.04113	0.02472	0.07068	0.01005
374	0.13301	0.02051	0.09294	0.01307	0.20633	0.01148	0.09025	0.01449
427	0.08127	0.01698	0.02177	0.01280	0.03283	0.02546	0.05820	0.02711
436	0.37545	0.01157	0.04606	0.01765	0.13370	0.01979	0.03486	0.01970
465	0.13109	0.01960	0.03856	0.03205	0.15184	0.02463	0.02737	0.03445
466	0.09261	0.04477	0.03367	0.03384	0.27403	0.02135	0.19652	0.02791
553	0.10078	0.00848	0.03222	0.01348	0.00064	0.01112	0.00524	0.01914

Πίνακας 0.6: Μέσες τιμές στις οποίες συγκλίνουν τα χαρακτηριστικά οδηγικής συμπεριφοράς (Ταξίδια μέσης χρονικής διάρκειας 20 min.)

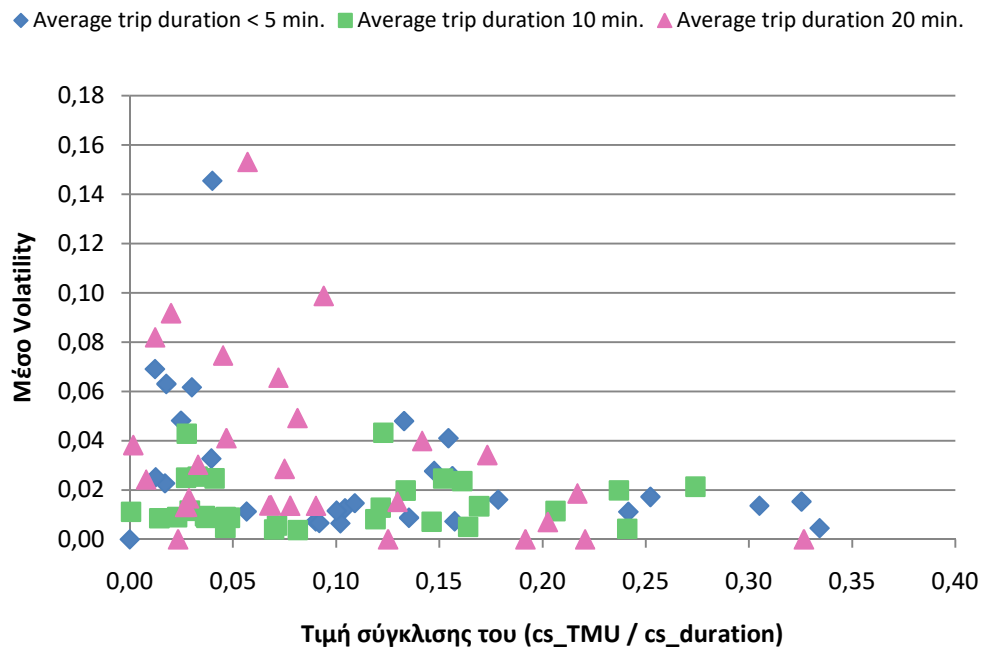
USER_ID	DUR 20							
	HARSH ACCELERATION		HARSH BRAKING		TIME MOBILE USAGE		SUM SPEEDING	
	cs_ha / totaldist	volatility	cs_hb / totaldist	volatility	cs_TMU / duration	volatility	cs_sum_speeding / duration	volatility
3	0.08560	0.03453	0.03994	0.06163	0.01221	0.08195	0.08098	0.02739
4	0.18717	0.00594	0.09540	0.00635	0.12971	0.01522	0.25495	0.00534
9	0.36664	0.00449	0.06457	0.00777	0.02678	0.01310	0.09281	0.00911
10	0.17239	0.02420	0.04607	0.02345	0.09394	0.09872	0.13004	0.03415
11	0.21679	0.00561	0.08261	0.00932	0.09017	0.01360	0.30058	0.00481
18	0.23148	0.00369	0.17917	0.00371	0.20256	0.00699	0.07226	0.00694
19	0.08420	0.01447	0.04643	0.02126	0.00786	0.02422	0.04441	0.01394
25	0.22436	0.01789	0.05564	0.03245	0.21692	0.01860	0.15789	0.02512
35	0.10794	0.05063	0.04749	0.15179	0.04517	0.07457	0.09403	0.05432
46	0.15549	0.01837	0.05107	0.02051	0.03303	0.03024	0.09733	0.04476
94	0.13109	0.01912	0.05699	0.02451	0.08121	0.04917	0.10372	0.02135
119	0.02831	0.02766	0.01618	0.02770	0.14165	0.03990	0.24706	0.03850
124	0.16358		0.05077		0.19170		0.08911	
132	0.08336	0.04397	0.01667	0.06905	0.07198	0.06559	0.04693	0.07167
140	0.04595	0.01744	0.02142	0.01606	0.02861	0.01662	0.09972	0.01212
143	0.16035	0.00608	0.06607	0.01143	0.06835	0.01385	0.06098	0.01115
146	0.19853	0.00746	0.06984	0.00885	0.07773	0.01362	0.06948	0.01158
147	0.11005	0.02036	0.06396	0.02832	0.17321	0.03420	0.10500	0.02124
149	0.09159		0.06297		0.02332		0.31705	
154	0.00757	0.02250	0.00142	0.04545	0.01995	0.09171	0.05280	0.01657
159	0.15611	0.00732	0.06314	0.00843	0.02777	0.01325	0.03221	0.01760
257	0.14060	0.00720	0.05999	0.00780	0.06756	0.01372	0.12786	0.00683
332	0.11783	0.01931	0.04305	0.01819	0.04678	0.04105	0.12652	0.01729
374	0.05615		0.02808		0.32667		0.13800	
427	0.08875	0.03257	0.03451	0.07896	0.05701	0.15308	0.13070	0.04586
436	0.43664		0.02239		0.12515		0.01346	
465	0.09357	0.02956	0.04954	0.06377	0.07498	0.02863	0.02100	0.08406
466	0.08246		0.01374		0.22064		0.42783	
553	0.05032		0.01677		0.00164	0.03821	0.00556	



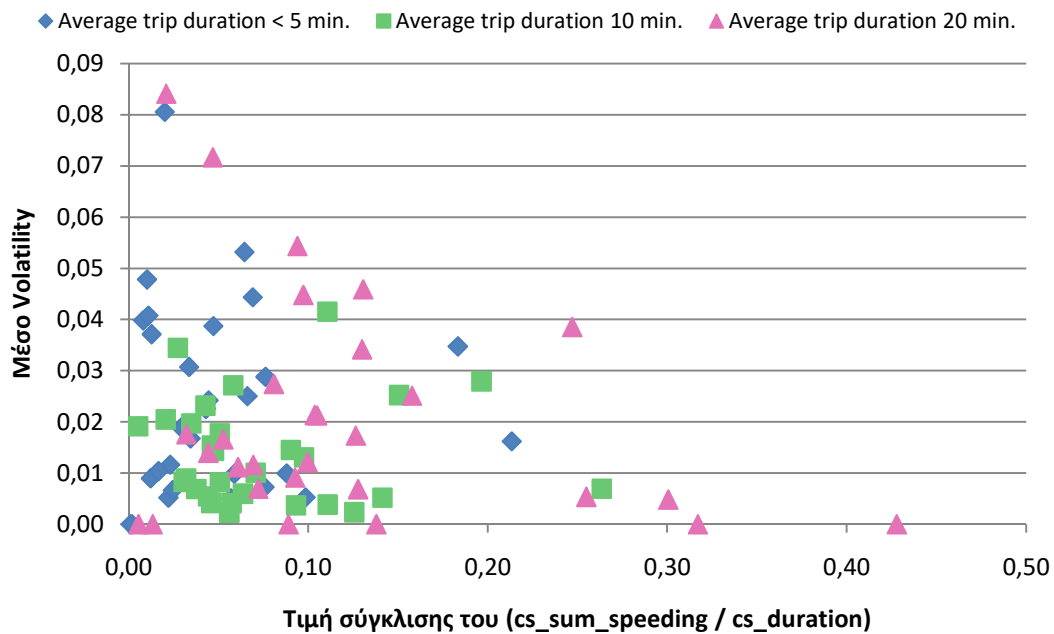
Διάγραμμα 0.5: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς τον αριθμό απότομων επιταχύνσεων ανά χιλιόμετρο



Διάγραμμα 0.6: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς τον αριθμό απότομων φρεναρισμάτων ανά χιλιόμετρο



Διάγραμμα 0.7: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς το ποσοστό του χρόνου χρήσης κινητού τηλεφώνου κατά την οδήγηση



Διάγραμμα 0.8: Επιθετικότητα - Μεταβλητότητα οδηγών ως προς το ποσοστό του χρόνου υπέρβασης του ορίου ταχύτητας κατά την οδήγηση

Ακολουθούν κάποια διαγράμματα σύγκλισης τυχαίων οδηγών, ενδεικτικά των διαφορετικών περιπτώσεων που υπάρχουν. Έχουν επισημανθεί τα σημεία από τα οποία και μετά η ανάλυση έδειξε ότι υπάρχει σύγκλιση, τόσο για τα μεγέθη σωρευτικών αθροισμάτων, όσο και για μέγεθος της μεταβλητότητάς τους. Για λόγους εποπτείας και σύγκρισης, παρατίθενεται στο ίδιο διάγραμμα, με χρήση διαφορετικών κατακόρυφων αξόνων.

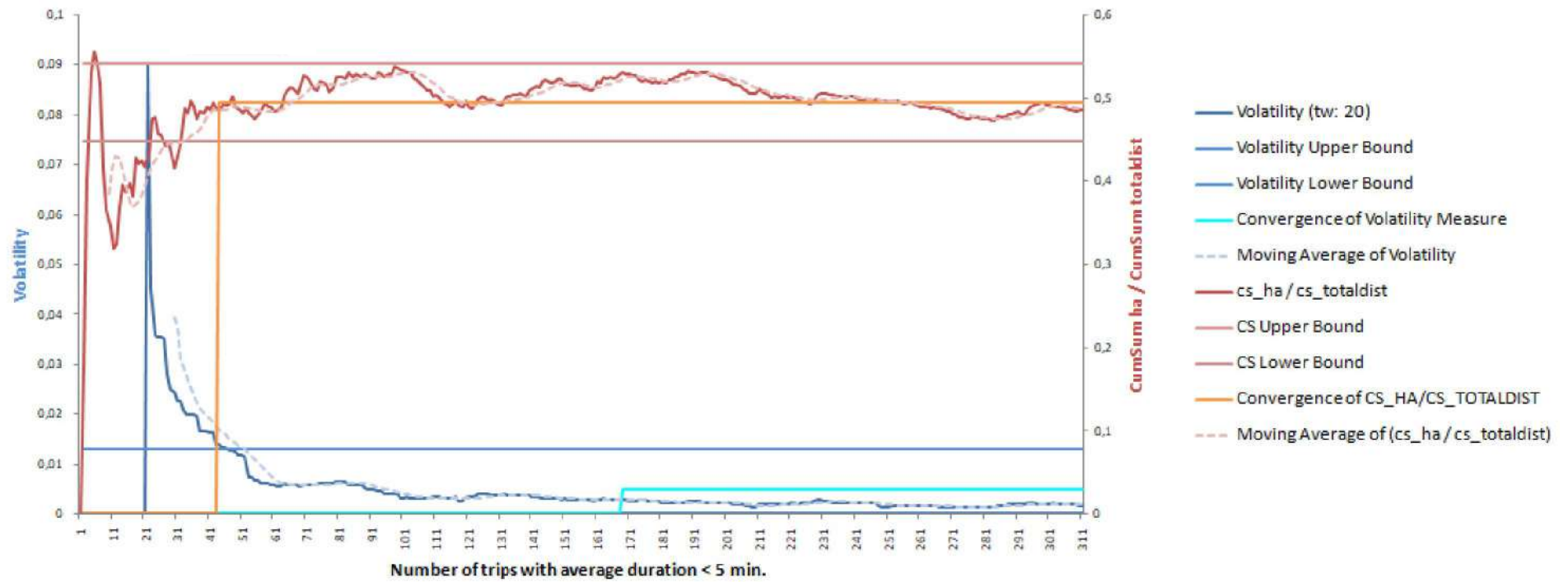
Το πρώτο διάγραμμα αφορά στον οδηγό «9» για ταξίδια μέσης διάρκειας < 5 λεπτών, και παρουσιάζει τη μεταβολή των συμβάντων απότομων επιταχύνσεων ανά χιλιόμετρο. Ο ρυθμός σύγκλισης του μεγέθους $cs_ha / cs_totaldist$ είναι ταχύτερος από αυτόν του Volatility του ίδιου μεγέθους. Πρόκειται για έναν «επιθετικό» οδηγό ως προς τις απότομες επιταχύνσεις, αφού η μέση τιμή στην οποία συγκλίνει είναι 0,4851 ha/km, αλλά με σταθερή (non-volatile) συμπεριφορά, αφού η μέση τιμή στην οποία συγκλίνει το Volatility του ίδιου μεγέθους είναι 0,0049.

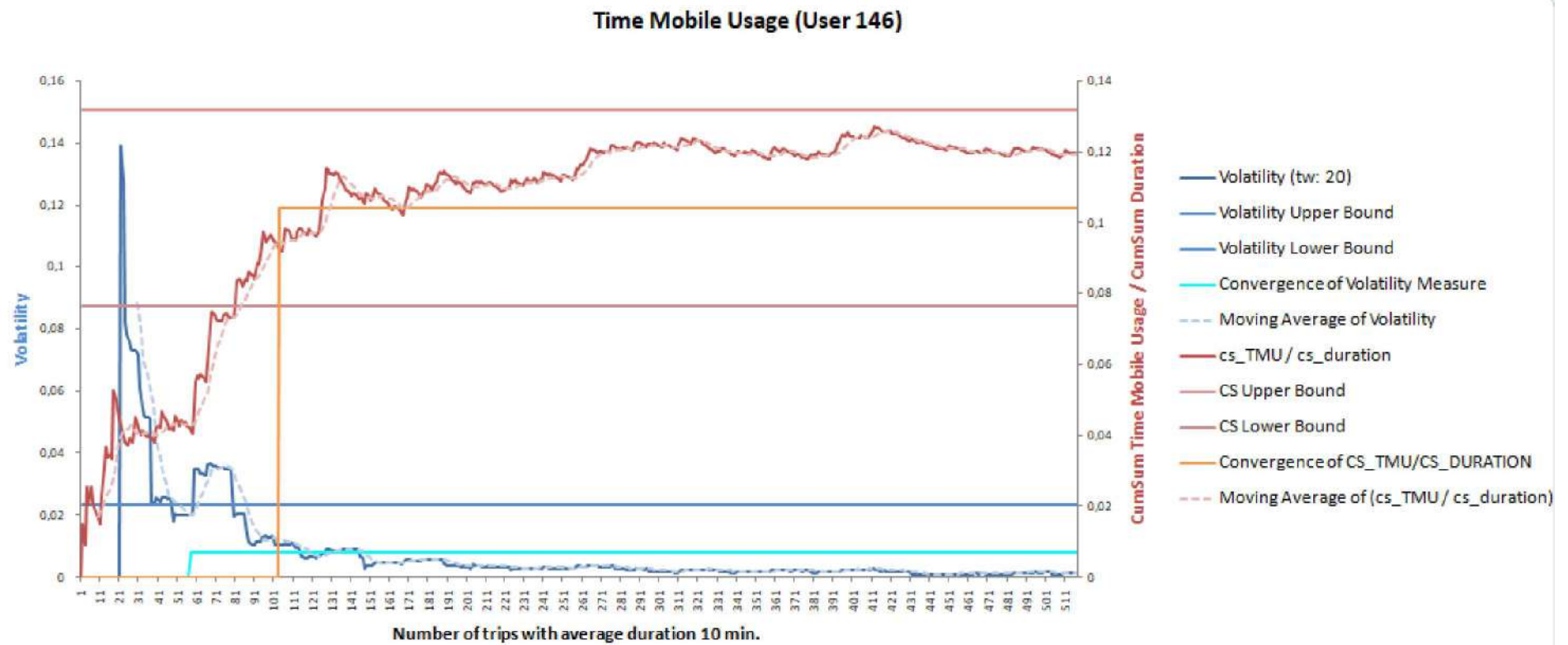
Το δεύτερο διάγραμμα αφορά στον οδηγό «146» για ταξίδια μέσης διάρκειας 10 λεπτών, και παρουσιάζει τη μεταβολή του ποσοστού χρόνου χρήσης κινητού τηλεφώνου κατά τη διάρκεια της οδήγησης. Ο ρυθμός σύγκλισης του μεγέθους $cs_TMU / cs_duration$ είναι πιο αργός από αυτόν του Volatility του ίδιου μεγέθους.

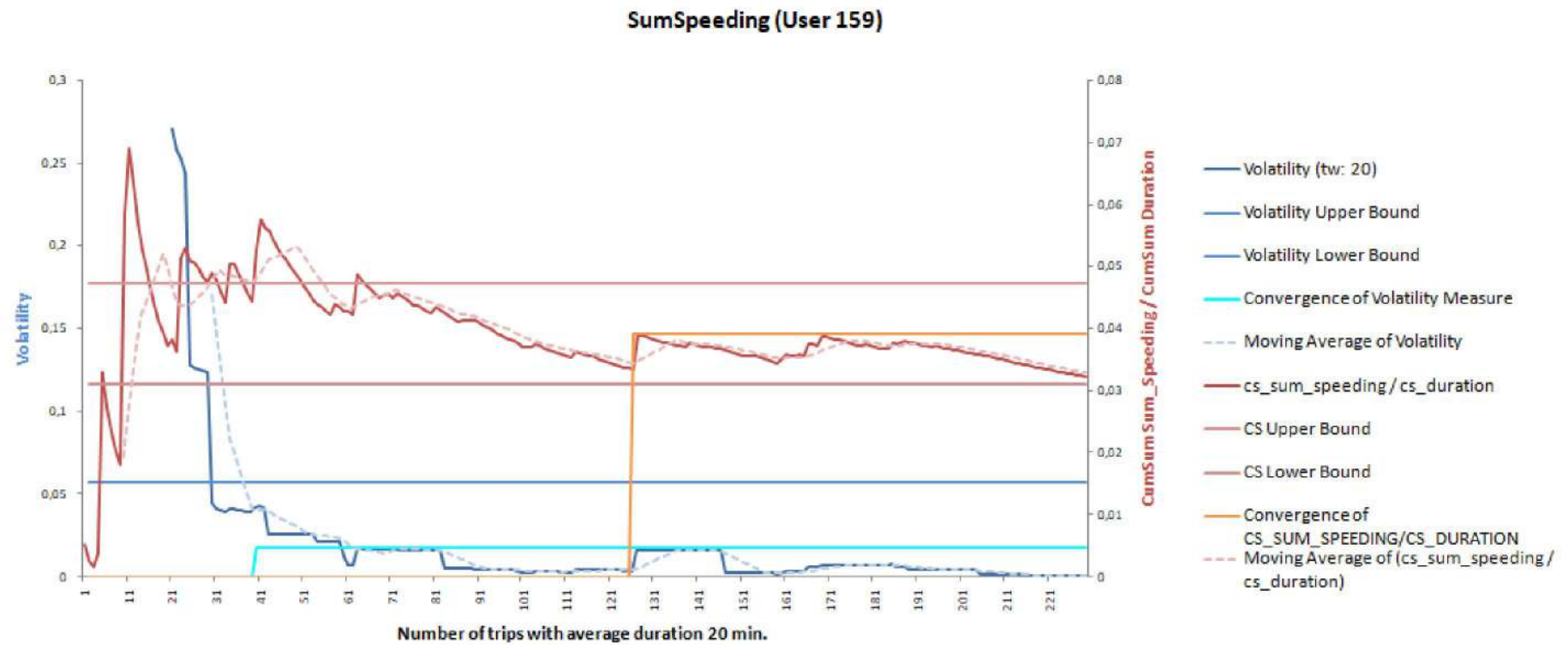
Το επόμενο διάγραμμα αφορά στον οδηγό «159» για ταξίδια μέσης διάρκειας 20 λεπτών, και παρουσιάζει τη μεταβολή του ποσοστού χρόνου υπέρβασης του ορίου ταχύτητας κατά τη διάρκεια της οδήγησης. Ο ρυθμός σύγκλισης του μεγέθους $cs_TMU / cs_duration$ είναι πολύ πιο αργός από αυτόν του Volatility του ίδιου μεγέθους.

Το τελευταίο διάγραμμα αφορά στον οδηγό «94» για ταξίδια μέσης διάρκειας 10 λεπτών, και παρουσιάζει τη μεταβολή των συμβάντων απότομων επιβραδύνσεων ανά χιλιόμετρο. Τα δύο μεγέθη ($cs_hb / cs_totaldist$ και Volatility του ίδιου μεγέθους) έχουν ίδιο ρυθμό σύγκλισης.

Harsh Acceleration (User 9)







Harsh Braking (User 94)

