



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής

**Χωρική Ανάλυση Οδηγικής Συμπεριφοράς
σε Αστικά Οδικά Δίκτυα**

Διπλωματική Εργασία

Γεωργία Λαγού

Επιβλέπουσα Καθηγήτρια : Ελένη Ι. Βλαχογιάννη
Επίκουρος Καθηγήτρια Σχολής Πολιτικών Μηχανικών ΕΜΠ

Αθήνα, Ιούλιος 2018

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας, ολοκληρώνεται και ο κύκλος των προπτυχιακών μου σπουδών στη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου. Με την ευκαιρία αυτή, θα ήθελα να ευχαριστήσω θερμά όλους όσους συνέβαλαν στη διαδικασία αυτή.

Αρχικά, θα ήθελα να ευχαριστήσω την επιβλέπουσα καθηγήτρια μου κυρία Ελένη Βλαχογιάννη για το ενδιαφέρον και την όρεξη που έδειξε καθ' όλη τη διάρκεια της συνεργασίας μας, με παρατηρήσεις και συμβουλές σε όλα τα στάδια εκπόνησης της εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω όλους μου τους φίλους, νέους και παλιούς, που στάθηκαν στο πλάι μου σε αυτήν την προσπάθεια με τον δικό τους τρόπο ο καθένας γεμίζοντας με αναμνήσεις από αυτά τα χρόνια. Εύχομαι να ζήσουμε πολλές ακόμα αξέχαστες στιγμές μαζί υπερπηδώντας κάθε εμπόδιο.

Τέλος, τίποτα δεν θα ήταν εφικτό χωρίς την συμπαράσταση της οικογένειας μου. Ένα μεγάλο ευχαριστώ στους γονείς μου Νίκο και Ευγενία, και τα αδέρφια μου Μάρθα και Γιάννη για την έμπρακτη στήριξή τους σε όλα τα σχολικά και φοιτητικά μου χρόνια. Ήταν οι άνθρωποι που με την αγάπη τους, την υπομονή τους και την διαρκή υποστήριξή τους έκαναν τις δύσκολες στιγμές να μοιάζουν ευκολότερες.

Χωρική Ανάλυση Οδηγικής Συμπεριφοράς

σε Αστικά Οδικά Δίκτυα

Συγγραφέας Διπλωματικής Εργασίας: Γεωργία Λαγού

Επιβλέπουσα Καθηγήτρια: Ελένη Ι. Βλαχογιάννη

ΣΥΝΟΨΗ

Η αυξημένη αστικοποίηση και η εκτεταμένη χρήση των Ι.Χ. αυτοκινήτων αυξάνει την ανάγκη για μετακίνηση και καθιστά σημαντική την βελτίωση του επιπέδου της οδηγικής συμπεριφοράς. Σημαντικό εργαλείο προς την κατεύθυνση αυτή, είναι η χωρική απεικόνιση και ανάλυση των οδηγικών συμβάντων σε χάρτες επικινδυνότητας του οδικού δικτύου. Προς εξυπηρέτηση του σκοπού αυτού, η έρευνα χρησιμοποιεί πρωτογενή δεδομένα παρακολουθούμενων μεγεθών από έξυπνα κινητά τηλέφωνα ικανών να περιγράψουν την οδηγική συμπεριφορά. Τα δεδομένα αυτά αποτελούν στοιχεία εισόδου για τους αλγορίθμους ομαδοποίησης της οδηγικής συμπεριφοράς που αναπτύχθηκαν. Βάσει των αποτελεσμάτων από τους παραπάνω αλγορίθμους, εντοπίστηκαν τα επικίνδυνα σημεία του δικτύου και αξιολογήθηκε η ομοιομορφία στην οδηγική συμπεριφορά ανάμεσα σε διαφορετικούς χρήστες της εφαρμογής. Τέλος, έγινε η χωρική απεικόνιση των ομάδων που προέκυψαν από την εφαρμογή των αλγορίθμων σε χάρτες επικινδυνότητας με σκοπό να αποτελέσουν εργαλείο ενημέρωσης και προειδοποίησης των οδηγών του δικτύου.

Λέξεις κλειδιά: Οδηγική Συμπεριφορά, Πληθοπορισμός, Επικινδυνότητα, Ομαδοποίηση Συμπεριφοράς, Χωρική Ανάλυση, Χωρική Απεικόνιση, DBSCAN.

Spatial Analysis of Driving Behavior

in Urban Road Networks

Thesis Author: Georgia Lagou

Supervising Professor: Eleni I. Vlahogianni

ABSTRACT

Increased urbanization and widespread use of private vehicles increases the need for mobility and makes it important to improve the level of driving behavior. An important tool in this direction is the spatial mapping and analysis of driving incidents on road hazard maps. To serve this purpose, research uses raw data of tracked sizes from smartphones capable of describing driving behavior. These data are input elements for the clustering algorithms of the driving behavior developed. Based on the results from the above algorithms, the network's dangerous points were identified and uniformity in driving behavior among different users of the application was assessed. Finally, the spatial representation of the groups resulting from the application of the algorithms on risk maps was made to serve as a tool for informing and warning the drivers of the network.

Keywords: Driving Behaviour, Crowdsourcing, Risky Driving, Clustering Analysis, Spatial Analysis, Spatial Visualization, DBSCAN.

ΠΕΡΙΛΗΨΗ

Η εξέλιξη της τεχνολογίας σε συνδυασμό με τις οικονομικές δυνατότητας εντείνουν το φαινόμενο της αστικοποίησης και κατά συνέπεια αυξάνουν τις ανάγκες των ανθρώπων για μετακίνηση. Οι μετακινήσεις αποτελούν αντικείμενο μείζονος σημασίας για τη ζωή των ανθρώπων. Για το λόγο αυτό είναι απαραίτητη η σωστή διαχείριση του οδικού συστήματος, για βελτίωση της οδηγικής συμπεριφοράς προκειμένου οι μετακινήσεις να γίνουν αποδοτικότερες, με σεβασμό στο περιβάλλον, να συμβάλλουν με το δικό τους τρόπο στην ανάπτυξη της οικονομίας και να παρέχουν ασφάλεια στους μετακινούμενους.

Σκοπός της παρούσας διπλωματικής εργασίας είναι να γίνει χωρική απεικόνιση των ακραίων οδηγικών συμπεριφορών με στόχο την πληροφόρηση και την έγκαιρη προειδοποίηση των οδηγών. Επιπλέον, κρίνεται χρήσιμη η ομαδοποίηση των οδηγικών συμπεριφορών των διαφόρων χρηστών προκειμένου να αξιολογηθεί η ομοιομορφία που τυχόν παρουσιάζουν στη λήψη αποφάσεων κατά την οδήγηση. Από την ανασκόπηση στη βιβλιογραφία, κρίθηκε σκόπιμο η προσέγγιση του ζητήματος να γίνει με αλγορίθμους ομαδοποίησης και έπειτα με απεικόνιση των αποτελεσμάτων σε χάρτες επικινδυνότητας.

Η έρευνα που ακολούθησε βασίστηκε σε βάσεις δεδομένων από έξυπνα κινητά τηλέφωνα μέσω της εφαρμογής OSeven. Οι βάσεις δεδομένων που χρησιμοποιήθηκαν αφορούν ταξίδια ενός μεμονωμένου χρήστη και πολλών τυχαίων χρηστών εντός του οδικού δικτύου της Αθήνας. Οι βάσεις επιλέχθηκαν με σκοπό να γίνει η σύγκριση των συμπεριφορών των οδηγών και να αξιολογηθεί η ομοιομορφία που παρουσιάζουν κατά τη διάρκεια της οδήγησης. Τα οδηγικά συμβάντα τα οποία εξετάστηκαν ήταν το σύνολο των απότομων επιταχύνσεων, το σύνολο των απότομων επιβραδύνσεων και το ποσοστό του χρόνου του ταξιδιού όπου πραγματοποιήθηκε υπέρβαση του ορίου ταχύτητας. Η προσέγγιση αυτή σχετίζεται με την επιθετικότητα και το βαθμό επικινδυνότητας της οδήγησης.

Στην επόμενη φάση, ακολούθησε ο προσδιορισμός των απαραίτητων στοιχείων που χρειάστηκαν για εισαγωγή στους αλγορίθμους ομαδοποίησης. Αναλυτικότερα, έγινε διάκριση των συμβάντων ανάλογα με το αν αυτά πραγματοποιούνται σε μια συγκεκριμένη χρονική στιγμή ή αν έχουν διάρκεια μέσα στο ταξίδι. Ακολούθησε ο προσδιορισμός των κατάλληλων παραμέτρων ξεχωριστά για κάθε αλγόριθμο και για κάθε βάση δεδομένων. Με τα αποτελέσματα των αλγορίθμων, δηλαδή με τις ομάδες οδηγικών συμπεριφορών που δημιουργήθηκαν, γίνεται η χωρική απεικόνισή τους σε χάρτες επικινδυνότητας. Ως εκ τούτου, διακρίνεται η ομοιομορφία στη συμπεριφορά του ενός μεμονωμένου οδηγού και των πολλών τυχαίων χρηστών. Επιπλέον, αναλύονται τα πιο επικίνδυνα σημεία του οδικού δικτύου της Αθήνας, σημεία δηλαδή τα οποία εμφανίζονται και στις δύο περιπτώσεις οδηγών.

Η ανάπτυξη των αλγορίθμων εκπλήρωσε τους σκοπούς της παρούσας διπλωματικής εργασίας. Δημιουργήθηκε μια διαδικασία ομαδοποίησης των οδηγικών συμπεριφορών και απεικόνισης αυτών σε χάρτες επικινδυνότητας του οδικού δικτύου. Εντοπίστηκαν τα επικίνδυνα σημεία του δικτύου και διερευνήθηκαν οι λόγοι για τους οποίους κρίθηκαν αυτά ως επικίνδυνα. Η διαδικασία που παρουσιάζεται μπορεί να συνεισφέρει αφενός σε ατομικό επίπεδο στον κάθε χρήστη ξεχωριστά με στόχο κυρίως την πληροφόρηση του για λήψη σωστών αποφάσεων κατά τη διάρκεια της οδήγησης αφετέρου σε επίπεδο δικτύου για βελτίωση του συστήματος διαχείρισης της οδηγικής συμπεριφοράς και της οδικής ασφάλειας.

Κατά τη διάρκεια της έρευνας τέθηκαν περιορισμοί και αναπτύχθηκαν και επιπλέον προβληματισμοί. Αυτά γεννούν σημεία στα οποία κρίνεται απαραίτητη η περαιτέρω εμβάθυνση. Αρχικά, είναι χρήσιμη η εισαγωγή επιπλέον παραγόντων που επηρεάζουν την οδηγική συμπεριφορά, όπως η γεωμετρία της οδού ή οι κυκλοφοριακές συνθήκες, προκειμένου να συγκριθούν οι χάρτες επικινδυνότητας και να αξιολογηθούν περαιτέρω λόγοι που επηρεάζουν την επιθετικότητα στην οδήγηση. Τέλος, κρίνεται αναγκαία η εφαρμογή περισσότερων αλγορίθμων ομαδοποίησης, ειδικότερα για τα μη σημειακά συμβάντα, προκειμένου να συγκριθούν τα αποτελέσματα των δύο μεθόδων αλλά και να γίνει μια βελτιστοποίηση της διαδικασίας που παρουσιάζεται στα πλαίσια της παρούσας διπλωματικής εργασίας.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ	i
ΣΥΝΟΨΗ	ii
ABSTRACT	iii
ΠΕΡΙΛΗΨΗ	iv
ΠΕΡΙΕΧΟΜΕΝΑ.....	vi
ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ.....	viii
ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ	x
ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ	xi
1. ΕΙΣΑΓΩΓΗ	1
1.1 Η Ανάγκη για Βιώσιμη Κινητικότητα	1
1.2 Οδηγικά Συμβάντα	4
1.3 Η Ανθρώπινη Συμπεριφορά κατά την Οδήγηση.....	5
1.4 Σκοπός Διπλωματικής.....	6
1.5 Διάρθρωση της Διπλωματικής	7
2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	9
2.1 Οδηγική Συμπεριφορά	9
2.1.1 Πληθοπορισμός και Έξυπνα Κινητά Τηλέφωνα	10
2.1.2 Μεγέθη Αξιολόγησης Οδηγικής Συμπεριφοράς	11
2.1.3 Παράγοντες Οδηγικής Συμπεριφοράς	12
2.2 Επικινδυνότητα και Οδική Ασφάλεια	14
2.2.1 Δείκτες Επικινδυνότητας.....	14
2.2.2 Παράγοντες Επικινδυνότητας	16
2.2.3 Είδη Ερευνών.....	16
2.2.4 Εργαλείο Επισήμανσης Επικίνδυνων Σημείων.....	18
2.2.5 Χάρτες Οδικής Ασφάλειας	20
2.3 Ομαδοποίηση Συμπεριφοράς.....	23
2.4 Συμπεράσματα Βιβλιογραφίας	27
3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ	29
3.1 Προσέγγιση Και Διάγραμμα Ροής Εργασιών	29
3.2 Θεωρητικό Υπόβαθρο	31
3.2.1 Ομαδοποίηση.....	31
3.2.2 Αλγόριθμος DBSCAN.....	34
3.2.3 Αλγόριθμος Εξόρυξης Προτύπων Ακολουθιών	38
4. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	41

4.1 Συλλογή Και Επεξεργασία Δεδομένων.....	41
4.2 Ανάλυση Απότομων Επιταχύνσεων Και Απότομων Επιβραδύνσεων	46
4.2.1 Προκαταρκτική Επεξεργασία Δεδομένων	46
4.2.2 Ανάλυση Συμπεριφοράς Μεμονωμένου Χρήστη.....	46
4.2.3 Ανάλυση Συμπεριφοράς Πολλών Χρηστών.....	49
4.3 Ανάλυση Συμβάντων Υπερβασης Ορίου Ταχύτητας	57
4.3.1 Προκαταρκτική Επεξεργασία Δεδομένων	57
5. ΣΥΜΠΕΡΑΣΜΑΤΑ	63
5.1 Εισαγωγή	63
5.2 Βασικά Συμπεράσματα.....	64
5.3 Προεκτάσεις Έρευνας.....	66
6. ΒΙΒΛΙΟΓΡΑΦΙΑ	69

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1. 1: Διαχρονική μεταβολή στην κατανομή των μετακινήσεων κατά μέσο (προσαρμογή από (Vuchic, 2007)).	2
Εικόνα 1. 2: Διαχρονική χρήση του ΙΧ αυτοκινήτου (1990-2009) (Goodwin, 2012).	2
Εικόνα 1. 3: Βασικοί πυλώνες για εξασφάλιση βιώσιμης κινητικότητας.	3
Εικόνα 2. 1: Περιβάλλον εφαρμογής (Πηγή:(Weekley et al., 2016)).	19
Εικόνα 2. 2: Χάρτης Επικινδυνότητα και Χάρτης Οδικής Ασφάλειας (Πηγή:(Dang, Kubo, Sato, & Namatame, 2015)).	20
Εικόνα 2. 3: Χάρτης Οδικής Ασφάλειας-Περιοχή Υψηλού Κινδύνου (Πηγή:(Dang et al., 2015)).	21
Εικόνα 2. 4:Χάρτης Οδικής Ασφάλειας-Περιοχή Χαμηλού Κινδύνου (Πηγή:(Dang et al., 2015)).	22
Εικόνα 2. 5: Κατανομή συγκρούσεων.	24
Εικόνα 2. 6: Αλγόριθμος K-means.	25
Εικόνα 2. 7: Αλγόριθμος DBSCAN.	25
Εικόνα 2. 8: Αλγόριθμος CB3N.	25
Εικόνα 2. 9: Αλγόριθμος DTH3N.	26
Εικόνα 3. 1: Γραφική απεικόνιση τριών ομάδων.	31
Εικόνα 3. 2: Σύνολο σημείων για ομαδοποίηση.	32
Εικόνα 3. 3: Τρόποι διαχωρισμού σημείων σε ομάδες.	32
Εικόνα 3. 4: Κατηγοριοποίηση σημείων.	34

Εικόνα 4. 1: Απεικόνιση του συνόλου των σημειακών συμβάντων μεμονωμένου χρήστη.	42
Εικόνα 4. 2: Απεικόνιση του συνόλου των συμβάντων υπέρβασης ορίου ταχύτητας μεμονωμένου χρήστη.	43
Εικόνα 4. 3: Απεικόνιση του συνόλου των σημειακών συμβάντων πολλών χρηστών.	44
Εικόνα 4. 4: Απεικόνιση του συνόλου των συμβάντων υπέρβασης ορίου ταχύτητας πολλών χρηστών.	45
Εικόνα 4. 5: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιταχύνσεων μεμονωμένου χρήστη.	48
Εικόνα 4. 6: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιβραδύνσεων μεμονωμένου χρήστη.	49
Εικόνα 4. 7: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιταχύνσεων πολλών χρηστών.	52
Εικόνα 4. 8: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιβραδύνσεων πολλών χρηστών.	53
Εικόνα 4. 9: Κρίσιμη ομάδα για συμβάντα απότομων επιταχύνσεων.	55
Εικόνα 4. 10: Κρίσιμος κόμβος για συμβάντα απότομων επιταχύνσεων.	55
Εικόνα 4. 11: Κρίσιμη ομάδα για συμβάντα απότομων επιβραδύνσεων.	56
Εικόνα 4. 12: Χάρτης επικινδυνότητας για συμβάντα υπέρβασης του ορίου ταχύτητας μεμονωμένου χρήστη.	59
Εικόνα 4. 13: Χάρτης επικινδυνότητας για συμβάντα υπέρβασης του ορίου ταχύτητας πολλών χρηστών.	60
Εικόνα 4. 14: Κρίσιμη ομάδα για συμβάντα υπέρβασης του ορίου ταχύτητας.	61

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 3. 1: Παράδειγμα μορφής δεδομένων για αλγόριθμο τύπου frequent sequence mining.....	39
Πίνακας 3. 2: Αλληλουχίες προϊόντων και συχνότητα εμφάνισης.....	40
Πίνακας 4. 1: Δεδομένα προς επεξεργασία.....	45
Πίνακας 4. 2: Δεδομένα προς εισαγωγή στον αλγόριθμο DBSCAN.....	46
Πίνακας 4. 3: Ομάδες επικινδυνότητας σε σχέση με τον αριθμό των συμβάντων.....	50
Πίνακας 4. 4: Επικίνδυνες ομάδες για συμβάντα απότομων επιταχύνσεων.....	54
Πίνακας 4. 5: Δεδομένα προς εισαγωγή στον αλγόριθμο frequent sequence mining.....	57

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 2. 1: Ορισμός "χρόνου για κρούση" (Time to Collision – TTC).....	14
Διάγραμμα 2. 2: Ορισμός και σχηματική απεικόνιση των δεικτών TET και TIT.	15
Διάγραμμα 2. 3: Συσχέτιση μεταξύ πλήθους ομάδων και ακτίνας.	24
Διάγραμμα 3. 1: Διάγραμμα ροής εργασιών διπλωματικής εργασίας,.....	30
Διάγραμμα 4. 1: Σχέση ομάδων επικινδυνότητας και αριθμού σημείων απότομων επιταχύνσεων κάθε ομάδας.	50
Διάγραμμα 4. 2: Σχέση ομάδων επικινδυνότητας και αριθμού σημείων απότομων επιβραδύνσεων κάθε ομάδας.	51

1.ΕΙΣΑΓΩΓΗ

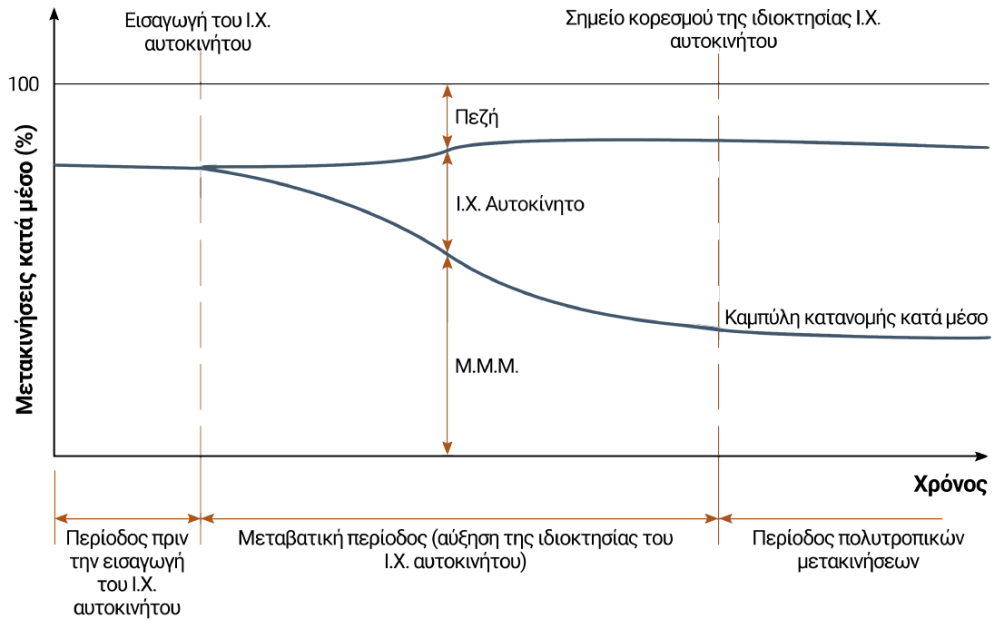
1.1 Η ΑΝΑΓΚΗ ΓΙΑ ΒΙΩΣΙΜΗ ΚΙΝΗΤΙΚΟΤΗΤΑ

Οι μεταφορές σήμερα αποτελούν ένα από τα βασικά μεγέθη και ταυτόχρονα μία από τις κύριες αιτίες των προβλημάτων που αντιμετωπίζει η σύγχρονη εποχή και η ζωή στον αστικό χώρο. Παράλληλα, αποτελούν σημαντικό παράγοντα της αστικής οικονομίας και αντικείμενο διαχείρισης, το οποίο σχετίζεται με αρκετούς τομείς όπως την αστική ζωή, την οικονομία ή την υγεία.

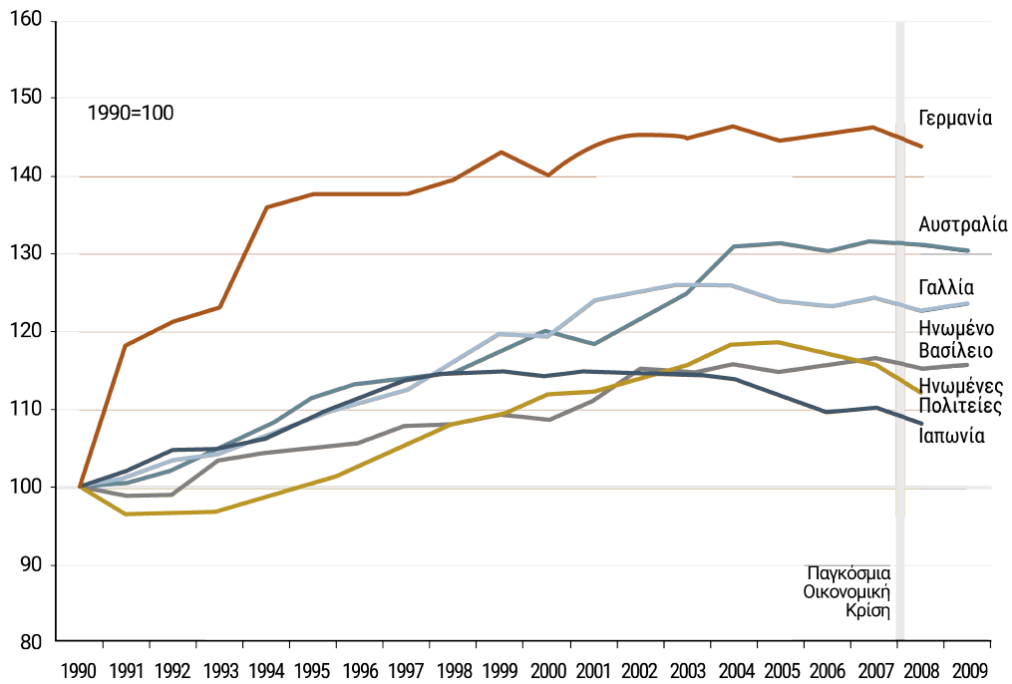
Η ραγδαία τεχνολογική ανάπτυξη του 20^{ου} αιώνα συνέβαλλε στην αύξηση της κινητικότητας μεγάλου μέρους του πληθυσμού. Έτσι, ο άνθρωπος έχει τη δυνατότητα να ικανοποιεί τις επιθυμίες του για προσπέλαση συγκεκριμένων προορισμών ή υπηρεσιών όλο και πιο εύκολα.

Στην Εικόνα 1. 1 φαίνεται η διαχρονική μεταβολή στην κατανομή των μετακινήσεων ανάμεσα στο ΙΧ αυτοκίνητο, τα ΜΜΜ και το περπάτημα με διάκριση στην περίοδο πριν και μετά την εισαγωγή του ΙΧ αυτοκινήτου. Παρατηρείται ότι με την εισαγωγή και τη σταδιακή αύξηση της χρήσης του αυτοκινήτου από τους ανθρώπους σημειώνεται μείωση τόσο των μετακινήσεων με τα ΜΜΜ όσο και των πεζή μετακινήσεων, μέχρι πρόσφατα όπου η ιδιοκτησία και κατά συνέπεια η χρήση του αυτοκινήτου έχει φθάσει στο ανώτατο σημείο (peak), σημείο κορεσμού, στον δυτικό κόσμο (Goodwin (2012)) (Εικόνα 1. 2). Σε αυτό συνηγορούν οι πολιτικές προώθησης των εναλλακτικών προς το ΙΧ αυτοκίνητο μέσων μεταφοράς καθώς και των συνδυασμένων μετακινήσεων που έχουν διαμορφωθεί τα τελευταία χρόνια. Η ισορροπία μεταξύ χρήσης του ΙΧ αυτοκινήτου και ΜΜΜ εξαρτάται από έναν αριθμό πολλών παραγόντων που αφορούν:

- Το επίπεδο εξυπηρέτησης των διαφόρων ΜΜΜ σε συνάρτηση με το κόστος τους.
- Τις υπάρχουσες πολιτικές σχετικά με την αστική ανάπτυξη (έλεγχος της αστικής ανάπτυξης ή ανεξέλεγκτη αστική διάχυση).
- Την προσφορά οδικής υποδομής και τις ισχύουσες πολιτικές στάθμευσης.
- Τις υπάρχουσες πολιτικές σχετικά με τον περιορισμό του ΙΧ αυτοκινήτου (Pucher, 2004).



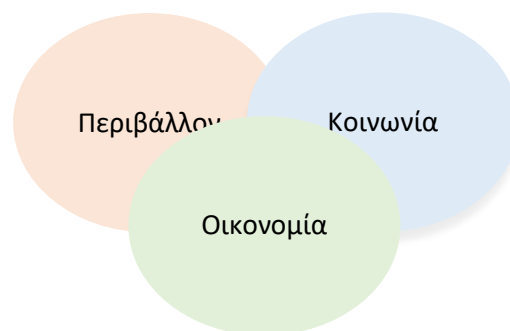
Εικόνα 1. 1: Διαχρονική μεταβολή στην κατανομή των μετακινήσεων κατά μέσο (προσαρμογή από (Vuchic, 2007)).



Εικόνα 1. 2: Διαχρονική χρήση του ΙΧ αυτοκινήτου (1990-2009) (Goodwin, 2012).

Τα μεταφορικά συστήματα ευθύνονται σε μεγάλο ποσοστό για τις εκπομπές διοξειδίου του άνθρακα και την παγκόσμια κατανάλωση ενέργειας. Για τον λόγο αυτό, οι τρέχουσες πρακτικές μεταφορών δεν κρίνονται βιώσιμες. Οι σύγχρονες πόλεις αντιμετωπίζουν προβλήματα αστικοποίησης και είναι εξαρτημένες από ιδιωτικής χρήσης οχήματα (Ι.Χ.). Η συνεχόμενη, λοιπόν, αύξηση της κινητικότητας αυτής οδηγεί σε ένταση των αρνητικών επιπτώσεων που σχετίζονται με την ατμοσφαιρική ρύπανση, τις κλιματολογικές αλλαγές, το θόρυβο, τον κυκλοφοριακό κορεσμό και τα ατυχήματα και κατά συνέπεια προκαλεί προβλήματα στη δημόσια υγεία.

Κρίνεται, λοιπόν απαραίτητη η ανάγκη για βιώσιμη κινητικότητα ως προϋπόθεση για διατήρηση του επιπέδου διαβίωσης κυρίως στις πόλεις. Για να είναι ένα σύστημα μεταφορών βιώσιμο, θα πρέπει να βελτιώνει το σύνολο των παραμέτρων του εξεταζόμενου συστήματος (Basbas & Politis, 2008). . Όπως περιγράφεται και σχηματικά στην Εικόνα 1. 3 η επίτευξη της βιωσιμότητας προϋποθέτει την ταυτόχρονη εξασφάλιση τριών βασικών πυλώνων (οικονομική, κοινωνική και περιβαλλοντική διάσταση) και συνεπώς και των υπο-συστατικών τους.



Εικόνα 1. 3: Βασικοί πυλώνες για εξασφάλιση βιώσιμης κινητικότητας.

1.2 ΟΔΗΓΙΚΑ ΣΥΜΒΑΝΤΑ

Μια ασυνήθιστη συμπεριφορά του οδηγού ή ένα είδος λάθους που πραγματοποιεί κατά τη διάρκεια της οδήγησης καλείται συμβάν (event). Τέτοιου είδους συμβάντα μπορεί να είναι απότομες επιταχύνσεις (harsh acceleration) ή επιβραδύνσεις (harsh break), χρονικές δηλαδή στιγμές στις οποίες ο οδηγός ασκεί περισσότερη δύναμη στο γκάζι ή στο φρένο αντίστοιχα. Επίσης, συμβάν θεωρείται και η οδήγηση με ταχύτητα μεγαλύτερη του ορίου ταχύτητας της εκάστοτε περιοχής. Τέτοιες συμπεριφορές από τους οδηγούς αποτελούν ένδειξη επιθετικής ή μη ασφαλούς οδήγησης.

Αρχικά, η παρακολούθηση και ο εντοπισμός τέτοιων ακραίων οδηγικών συμβάντων έχει σημαντική επιρροή στη ποιότητα του περιβάλλοντος. Όπως είναι γνωστό, μεγάλο ποσοστό της ατμοσφαιρικής ρύπανσης οφείλεται στις οδικές μεταφορές, εξ αιτίας του μεγάλου ποσοστού εκπομπών διοξειδίου του άνθρακα. Μέσω της επιθετικής οδήγησης η κατάσταση αυτή επιδεινώνεται. Η οδήγηση με σταθερό ρυθμό συνεπάγεται χαμηλότερη κατανάλωση καυσίμου. Αντίθετα, μια επιθετική οδήγηση έχει ως συνέπεια, εκτός των άλλων, πολύ μεγαλύτερη κατανάλωση καυσίμου. Η επιθετική αυτή συμπεριφορά, λοιπόν, οδηγεί στην εκπομπή περισσότερων ρύπων στην ατμόσφαιρα και συνεπώς δεν κρίνεται ως φιλική προς το περιβάλλον. Επιπλέον, η επιθετική οδήγηση αυξάνει σημαντικά τις δαπάνες συντήρησης του οχήματος λόγω του ότι φθείρεται ο μηχανολογικός εξοπλισμός του ή ότι δημιουργούνται υλικές ζημιές. Φυσικά δεν θα μπορούσε να παραλειφθεί το γεγονός ότι μια επιθετική συμπεριφορά οδηγεί στην αύξηση των οδηγικών ατυχημάτων.

Κρίνεται απαραίτητος ο εντοπισμός των ακραίων οδηγικών συμβάντων για τη μείωση της επιθετικής οδήγησης και όλων όσων αυτή προκαλεί, όπως αναφέρθηκαν και παραπάνω. Ο οδηγός είναι χρήσιμο να είναι σε θέση να αναγνωρίζει σημεία του οδικού δικτύου όπου παρατηρείται λανθασμένη οδήγηση από μεγάλο ποσοστό των οδηγών, προκειμένου να προειδοποιείται εγκαίρως και λαμβάνει πρόωρα τις κατάλληλες αποφάσεις.

1.3 Η ΑΝΘΡΩΠΙΝΗ ΣΥΜΠΕΡΙΦΟΡΑ ΚΑΤΑ ΤΗΝ ΟΔΗΓΗΣΗ

Οι αυξημένες δυνατότητες των σύγχρονων κινητών τηλεφώνων σε συνδυασμό με τον εύκολο προγραμματισμό τους, το μεγάλο ποσοστό διείσδυσής τους στην αγορά και την αποτελεσματική προσέγγισή τους από τρίτες εταιρείες λογισμικού με σκοπό την ανάπτυξη εφαρμογών, έχουν συμβάλλει στην ωρίμανσή τους και τη μετατροπή τους σε ένα αποτελεσματικό εργαλείο παρακολούθησης της ανθρώπινης συμπεριφοράς (Mitchell, 2009).

Με τον όρο «οδηγική συμπεριφορά», εννοούνται όλες οι φυσικές κινήσεις του οδηγού κατά τη διάρκεια της οδήγησης είτε ακόμα κατά τη διάρκεια ενός ταξιδιού στο οποίο βρίσκεται ως επιβάτης. Οι κινήσεις αυτές καταγράφονται από αισθητήρες έξυπνου κινητού τηλεφώνου. Τέτοιες κινήσεις μπορεί να είναι, για παράδειγμα, η χρήση κινητού τηλεφώνου ή η απότομη επιτάχυνση του οχήματος.

Η ανάλυση της οδηγικής συμπεριφοράς είναι χρήσιμη για τη βελτίωση της ασφάλειας και της κινητικότητας του κυκλοφοριακού συστήματος αλλά και σε άλλους ερευνητικούς τομείς. Για παράδειγμα, η παρακολούθηση της ανθρώπινης κινητικότητας μπορεί να ωφεληθεί άμεσα από τη δυνατότητα παρακολούθησης της συμπεριφοράς των ατόμων κατά τη μεταφορά τους από ένα σημείο σε ένα άλλο (Lazer et al., 2009). Αυτό με τη σειρά του θα επέτρεπε τη βελτίωση του πολεοδομικού σχεδιασμού (Zheng, Liu, Yuan, & Xie, 2011), την παρακολούθηση και αντιμετώπιση της εξάπλωσης των ασθενειών καθώς και άλλων πιθανών κινδύνων, καθώς και την παροχή πληροφοριών έκτακτης ανάγκης σχετικά με την ταχύτερη διαδρομή για να βοηθηθούν οι χαμένοι ή οι τραυματίες (Soper, 2012).

1.4 ΣΚΟΠΟΣ ΔΙΠΛΩΜΑΤΙΚΗΣ

Η παρούσα διπλωματική εργασία αποσκοπεί στην χωρική απεικόνιση της οδηγικής συμπεριφοράς και στην ομαδοποίηση αυτής για οδηγούς εντός του οδικού δικτύου της Αθήνας. Η παραγόμενη μεθοδολογία επιδιώκει την χωρική αναπαράσταση της συμπεριφοράς ενός δείγματος οδηγών. Με τον τρόπο αυτό, θα διαμορφωθούν χάρτες επικινδυνότητας μέσα στο οδικό δίκτυο, προκειμένου να οριστούν οι κρισιμότερες θέσεις του οδικού δικτύου για προειδοποίηση και έγκαιρη ενημέρωση των οδηγών.

Η έρευνα βασίζεται στην ανάπτυξη αλγορίθμων ομαδοποίησης της οδηγικής συμπεριφοράς διαφορετικών τυχαίων οδηγών προς βελτίωση της ικανότητας του χρήστη να λαμβάνει πληροφορίες για τη λήψη σημαντικών αποφάσεων κατά τη διάρκεια της οδήγησης.

Για την υλοποίηση των στόχων της έρευνας, αξιοποιούνται βάσεις δεδομένων διαφορετικών χρηστών εντός του οδικού δικτύου της Αθήνας της εφαρμογής έξυπνου κινητού τηλεφώνου OSeven. Μέσω της εφαρμογής αυτής παρακολουθούνται και καταγράφονται στοιχεία για τις μετακινήσεις των ατόμων που την χρησιμοποιούν. Η φύση τέτοιων δεδομένων, όπως αναλύεται και στη Βιβλιογραφική Ανασκόπηση που ακολουθεί, κρίνεται αναγκαία για την προσέγγιση τέτοιου είδους προβλημάτων οδηγικής συμπεριφοράς.

Οι αλγόριθμοι που χρησιμοποιούνται για την αντιμετώπιση του προβλήματος είναι:

- ο αλγόριθμος DBSCAN για ομαδοποίηση χωρικών δεδομένων για σημειακά οδηγικά συμβάντα, δηλαδή για απότομες επιταχύνσεις και επιβραδύνσεις.
- ένας αλγόριθμος εξόρυξης προτύπων ακολουθιών (frequent sequence mining) για την ομαδοποίηση χωροχρονικών δεδομένων, για συμβάντα δηλαδή υπέρβασης του ορίου ταχύτητας.

Η διαδικασία της ομαδοποίησης των δεδομένων έχει σκοπό τον εντοπισμό παρόμοιων οδηγικών συμπεριφορών εντός του οδικού δικτύου της Αθήνας από διαφορετικούς οδηγούς. Με τον τρόπο αυτό, είναι εύκολο να βρεθούν και να απεικονιστούν στο χώρο τα επικίνδυνα σημεία του οδικού δικτύου. Έτσι, ο κάθε οδηγός θα είναι σε θέση να αναγνωρίζει τα επικίνδυνα σημεία του δικτύου και να προειδοποιείται εγκαίρως για αυτά προκειμένου να λαμβάνει τις σωστές αποφάσεις κατά τη διάρκεια του ταξιδιού.

1.5 ΔΙΑΡΘΡΩΣΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ

Στην ενότητα αυτή, παρουσιάζεται η διάταξη των κεφαλαίων της διπλωματικής εργασίας, δίνοντας μια σύντομη περιγραφή για το αντικείμενο του καθενός.

Κεφάλαιο 1

Το παρόν πρώτο κεφάλαιο περιλαμβάνει τη παρουσίαση των προβλημάτων των μεταφορών της σύγχρονης εποχής και την αναγκαιότητα για στροφή στην βιώσιμη κινητικότητα. Εξηγεί, επίσης τη χρησιμότητα της ανάλυσης της οδηγικής συμπεριφοράς, μέσα από πληθοποριστικά δεδομένα. Τέλος, αναλύεται η έννοια της οδηγικής συμπεριφοράς κατά τη διάρκεια ενός ταξιδιού και προσδιορίζεται η έννοια των οδηγικών συμβάντων.

Κεφάλαιο 2

Στο δεύτερο κεφάλαιο πραγματοποιείται η βιβλιογραφική ανασκόπηση ερευνών που σχετίζονται με τη θεματολογία της παρούσας εργασίας. Αρχικά, εξετάζεται η ανάλυση της οδηγικής συμπεριφοράς μέσα από σύγχρονες πηγές πληροφοριών και από μεγέθη μέσα από τη βιβλιογραφία που την περιγράφουν με ακρίβεια. Έπειτα, γίνεται προσδιορισμός του όρου της επικινδυνότητας στην οδική συμπεριφορά και παρουσιάζονται τρόποι απεικόνισης της επικινδυνότητας αυτής.

Κεφάλαιο 3

Παρατίθεται το διάγραμμα ροής των εργασιών που ακολουθήθηκαν για την προσέγγιση του ζητήματος και παρουσιάζονται τα μεθοδολογικά εργαλεία που χρησιμοποιήθηκαν για την ανάλυση των δεδομένων.

Κεφάλαιο 4

Περιλαμβάνει τις διαδικασίες που χρησιμοποιήθηκαν για την χωρική αναπαράσταση των δεδομένων. Αρχικά, περιγράφονται οι βάσεις δεδομένων που χρησιμοποιούνται. Έπειτα, γίνεται σαφής προσδιορισμός των στοιχείων και των παραμέτρων που χρησιμοποιούνται στους αλγόριθμους. Το κεφάλαιο αυτό ολοκληρώνεται με την παρουσίαση των αποτελεσμάτων.

Κεφάλαιο 5

Στο Κεφάλαιο αυτό, γίνεται η παρουσίαση των συμπερασμάτων της έρευνας. Επιπλέον, παρατίθενται προτάσεις για περαιτέρω έρευνα εξαιτίας των περιορισμών που τέθηκαν αλλά και για προβληματισμούς για επίλυση νέων κατευθύνσεων.

Κεφάλαιο 6

Παρατίθεται η βιβλιογραφική ανασκόπηση που χρησιμοποιήθηκε για την καταγραφή της διπλωματικής εργασίας.

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Το κεφάλαιο αυτό περιλαμβάνει την ανάλυση των ερευνών σχετικών με τον τομέα της οδικής συμπεριφοράς, της επικινδυνότητας, των χωροχρονικών απεικονίσεων και την ομαδοποίηση της συμπεριφοράς των οδηγών. Στόχος της ανάλυσης αυτής είναι να εντοπιστούν τα κενά της έρευνας και να επιλεγεί η κατάλληλη μεθοδολογική προσέγγιση για αντιμετώπιση του αντικειμένου της διπλωματικής εργασίας.

2.1 ΟΔΗΓΙΚΗ ΣΥΜΠΕΡΙΦΟΡΑ

Η ραγδαία ανάπτυξη των πόλεων και η συνεπαγόμενη αλματώδης αύξηση των Ι.Χ. μόνο ως δυσμενής μπορεί να κριθεί για τα συστήματα μεταφορών, τόσο σε επίπεδο κινητικότητας όσο και περιβάλλοντος. Εξ' αιτίας της οικονομικής κρίσης, είναι αδύνατη η επένδυση σε μεγάλα έργα. Επομένως, κρίνεται αναγκαία η στροφή στην έρευνα της συμπεριφοράς του ίδιου του χρήστη σε απόπειρα βελτιστοποίησης της οδηγικής νοοτροπίας με απώτερο σκοπό την επίτευξη βιώσιμων συνθηκών στον τομέα των μεταφορών. Τα παραπάνω διασαφηνίζουν την ανάγκη παροχής δεδομένων ευρείας κλίμακας σε ερευνητές, που να σχετίζονται με τα χαρακτηριστικά των μετακινήσεων και μέσω επεξεργασίας των οποίων θα διαμορφώνονται συστάσεις προς τους χρήστες για αποδοτικότερο τρόπο οδήγησης.

Παραδοσιακά, κύρια πηγή δεδομένων για την ανάλυση της οδηγικής συμπεριφοράς αποτελούν έρευνες φυσικής οδήγησης, οι οποίες εμφανίζονται με τον όρο **Naturalistic Driving Studies (NDS)**. Μια συνηθισμένη έρευνα τέτοιου τύπου πραγματοποιείται με τη βοήθεια ειδικού εξοπλισμού, όπως κάμερες και GPS. Ωστόσο, μέσω των μετρήσεων αυτών, δεν είναι δυνατή η ανίχνευση όλων των συμβάντων που πραγματοποιούνται κατά τη διάρκεια της οδήγησης ούτε η αντίστοιχη επιθετικότητα των συμβάντων αυτών. Επιπλέον, τέτοιου είδους μετρήσεις κοστίζουν σε χρόνο και σε χρήμα για τους ερευνητές. Έτσι, εισάγεται μια νέα έννοια στη συλλογή δεδομένων.

2.1.1 Πληθοπορισμός και Έξυπνα Κινητά Τηλέφωνα

Μια εναλλακτική λύση για τη συλλογή δεδομένων που αφορούν συγκοινωνιακά προβλήματα είναι η άντληση αυτών μέσω έξυπνων κινητών τηλεφώνων (smartphones). Τα δεδομένα αυτά είναι γνωστά ως πληθοποριστικά δεδομένα (crowdsourced data). Με τον όρο πληθοπορισμός (crowdsourcing) όπως ορίζεται από τους (Estelles-Arolas & Gonzalez-Ladron-De-Guevara, 2012) εννοείται : μία μορφή συλλογικής διαδικτυακής δραστηριότητας στην οποία ένα άτομο, ένα ίδρυμα, ένας μη κερδοσκοπικός οργανισμός ή μία εταιρεία προτείνει σε μία ομάδα ατόμων με ποικίλες γνώσεις, ετερογένεια και αριθμό, μέσω μίας ανοικτής πρόσκλησης, να αναλάβουν εθελοντικά μια εργασία. Η ανάληψη της εργασίας, η οποία ποικίλλει σε πολυπλοκότητα και στο βαθμό στον οποίο είναι χωρισμένη και στην οποία το πλήθος πρέπει να συμμετάσχει με προσωπική εργασία, χρήματα, γνώση, εμπειρία, περιλαμβάνει πάντοτε αμοιβαίο όφελος και για τις δύο πλευρές.

Ειδικότερα, σε συγκοινωνιακά προβλήματα, ο όρος αυτός συναντάται ως **mobile crowdsourcing** και αφορά την εκτεταμένη χρήση έξυπνων κινητών τηλεφώνων ως μέσο άντλησης δεδομένων για την ανάλυση της οδηγικής συμπεριφοράς.

Η χρήση έξυπνων κινητών τηλεφώνων είναι ευρέως διαδεδομένη, καθώς παρουσιάζουν αρκετά πλεονεκτήματα τόσο λόγω της φύσης των χαρακτηριστικών τους όσο σε σύγκριση με άλλες προσεγγίσεις τηλεματικής, όπως (Calabrese et al., 2011; Toledo et al., 2008; Zhao, 2000):

- Υψηλή εισχώρηση ανάμεσα στο σύνολο των οδηγών.
- Ευκολία εκμετάλλευσης των λειτουργιών τους μέσω πληθώρας εφαρμογών.
- Χαμηλό κόστος ηλεκτρολογικού και μηχανολογικού τύπου αισθητήρες IMU (internal measurement units), όπως επιταχυνσιόμετρο, γυροσκόπιο και μαγνητόμετρο, αλλά και συστήματα GNSS (global navigational satellite system).
- Υπερτερούν σε επίπεδο εγκατάστασης και συντήρησης, ειδικά σε πειράματα μεγάλου όγκου οχημάτων.
- Διαρκής βελτίωση της αναλογίας απόδοσης-τιμής (Lane et al., 2010).
- Παροχή οπτικοακουστικών μέσω αλληλεπίδρασης με το χρήστη πριν (οργάνωση), κατά τη διάρκεια (γνωμοδότηση για τον τρόπο οδήγησης και επικέντρωση της προσοχής σε κρίσιμες καταστάσεις) και με το πέρας του ταξιδιού (αναφορά επίδοσης).
- Ευχερής και ασφαλής μεταφορά και διατήρηση δεδομένων σε online συστήματα (π.χ. Cloud) χάρη στην ασύρματη συνδεσιμότητα.
- Δεν έχουν ανάγκη σημαντικής επεξεργασίας όπως τα δεδομένα από GPS προς εξαγωγή στοιχείων για την οδηγική συμπεριφορά (Chowdhury, Chakravarty, & Balamuralidhar, 2014).

Από την άλλη, η βιβλιογραφία καταγράφει κάποια μειονεκτήματα της χρήσης τους (Handel et al., 2014):

- Απρόβλεπτη και αυθαίρετη τοποθέτηση στο όχημα
- Επίπεδα θορύβου δεδομένων για διαφορετικούς τύπους ταξιδιών με διαφορετικά μέσα μετακίνησης
- Άντληση μεγάλου μέρους της μπαταρίας της συσκευής.

Εξαιτίας των παραπάνω επηρεάζεται η ακρίβεια και η αξιοπιστία των δεδομένων που παρέχουν τα κινητά τηλέφωνα. Τα έξυπνα κινητά τηλέφωνα υπερεκτιμούν κρίσιμα περιστατικά και η απόδοση τους επηρεάζεται από τις οδικές συνθήκες που επικρατούν στο εκάστοτε σύστημα, ακόμα και από τη θέση του κινητού τηλεφώνου στο όχημα (Paefgen, Kehr, Zhai, & Michahelles, 2012).

Στη βιβλιογραφία υπάρχουν αναφορές για απόπειρες διόρθωσης των πρωτογενών δεδομένων από το επιταχυνσιόμετρο του κινητού. Οι (Vlahogianni & Barmounakis, 2017) ενσωματώνουν έναν αλγόριθμο επαναπροσδιορισμού συσκευής που εκμεταλλεύεται στοιχεία από γυροσκόπιο και GPS ώστε να διορθωθούν τα δεδομένα του έξυπνου τηλεφώνου. Με αυτόν τον τρόπο, αναπτύσσεται ένα πλαίσιο Μηχανικής Μάθησης βασισμένο σε rough set theory προς αναγνώριση κανόνων και μοτίβων οδηγικής συμπεριφοράς.

2.1.2 Μεγέθη Αξιολόγησης Οδηγικής Συμπεριφοράς

Από έρευνα των (Handel et al., 2014), ο προσδιορισμός των απαραίτητων μεγεθών (Figures of Merits-FoMs) για την αξιολόγηση της οδηγικής συμπεριφοράς επιτυγχάνεται σύμφωνα με τα παρακάτω κριτήρια:

- Παρατηρησιμότητα μεγέθους (FoM observability): Η συσχέτιση μεταξύ των μετρήσεων από τον αισθητήρα με το εξεταζόμενο μέγεθος.
- Στασιμότητα συμβάντος (Event stationarity): Το χρονικό μήκος κατά το οποίο τα γεγονότα που απαρτίζουν το μέγεθος καταγράφονται.
- Πραγματική συσχέτιση (Actual relevance): Η σημασία του μεγέθους για την αξιολόγηση ρίσκου από τον οδηγό.
- Επιρροή οδηγού (Driver's influence): Ο βαθμός που ο οδηγός επηρεάζει το εξεταζόμενο μέγεθος.

Επιπλέον, παρατίθεται μία σειρά από τα σημαντικότερα μεγέθη, που κατά την κρίση των ερευνητών περιγράφουν την οδηγική συμπεριφορά, τα οποία αξιολογούνται με βάση τα παραπάνω, εξαιρώντας ωστόσο την εκτίμηση των οδικών συνθηκών και την ένταση της υπάρχουσας κίνησης (Handel et al., 2014).

- Αριθμός απότομων επιταχύνσεων (Rapid acceleration)
- Αριθμός απότομων επιβραδύνσεων (Harsh braking)
- Υπέρβαση ορίου ταχύτητας (Speeding)
- Απαλότητα οδήγησης (Smoothness)
- Πλήθος απότομων ελιγμών (Swerving)
- Πλήθος στροφών υπό υψηλή ταχύτητα (Cornering)
- Οικολογικότητα οδήγησης (Eco-ness)
- Χρονική διάρκεια ταξιδιού (Elapsed time)
- Διανυόμενη απόσταση (Elapsed distance)
- Ώρα πραγματοποιούμενου ταξιδιού (Time of day)
- Τοποθεσία (Location)

Παρόλα αυτά τα μεγέθη, κρίσιμος παράγοντας που επηρεάζει καθοριστικά την απόσπαση του οδηγού κατά τη διάρκεια της οδήγησης θεωρείται η χρήση κινητού τηλεφώνου. Η αναγνώριση του συμβάντος αυτού συμβάλλει στην καλύτερη απεικόνιση της απόσπασης του οδηγού, αλλά και στον καθαρισμό του θορύβου προερχόμενου από απότομες κινήσεις της συσκευής. Οι αισθητήρες που ενεργοποιούνται επηρεάζουν τη διαδικασία αναγνώρισης των κρίσιμων καταστάσεων αλλά και των ορίων αναγνώρισης τους. Συνεπώς, με ορθή αναγνώριση της χρήσης κινητού τηλεφώνου μειώνονται τα false positive γεγονότα (συμβάντα που λανθασμένα λαμβάνονται ως κρίσιμα) και βελτιώνεται με αυτόν τον τρόπο η ικανότητα εντοπισμού της μεθόδου (Vlahogianni & Barmounakis, 2017).

2.1.3 Παράγοντες Οδηγικής Συμπεριφοράς

Σημαντικό πρόβλημα για την κατανόηση και την πρόβλεψη των προτύπων μετακίνησης αποτελεί η αναγνώριση των παραγόντων που οδηγούν στα πρότυπα οδηγικής συμπεριφοράς. Γενικά, για σχεδόν όλους τους τύπους οδηγικών συμπεριφορών, τρεις κατηγορίες παραγόντων είναι σημαντικές:

- Κοινωνικοδημογραφικοί παράγοντες
- Παράγοντες που σχετίζονται με τη δομή του περιβάλλοντος (πυκνότητα, ποικιλομορφία, σχεδιασμός και επίδραση απόστασης (Liu, Kang, & Wang, 2014))
- Παράγοντες που σχετίζονται με το ταξίδι (σκοπός ταξιδιού, χρόνος ταξιδιού, κόστος και αξιοπιστία).

Νέες έρευνες σχετικά με το πώς τα πρότυπα κινητικότητας μπορούν να ποικίλουν ανάλογα με το περιβάλλον είναι χρήσιμες για να γίνουν κατανοητές οι στρατηγικές αναζήτησης των ανθρώπων και πως αυτές μπορούν να αλλάξουν υπό διαφορετικές συνθήκες. Στην πραγματικότητα, έπειτα από έρευνα σε μεγάλα δεδομένα (Kang, Ma, Tong, & Liu, 2012) η εύρεση διαφορετικής κινητικότητας υποδηλώνει ότι η ανθρώπινη κίνηση δεν είναι απλώς αποτέλεσμα μιας μηχανικής διαδικασίας, αλλά επηρεάζεται από διάφορους παράγοντες που ποικίλουν από τόπο σε τόπο και από πληθυσμό σε πληθυσμό.

Έρευνες σε μεγάλες βάσεις δεδομένων μπορούν να βοηθήσουν στον εντοπισμό πρόσθετων παραγόντων που εξηγούν τις αποφάσεις και αναδεικνύουν το μηχανισμό λήψης αποφάσεων. Για παράδειγμα, αρκετές έρευνες δείχνουν τη σχέση μεταξύ των κοινωνικών δικτύων και της οδηγικής συμπεριφοράς στο χώρο και στο χρόνο (Eagle, Pentland, & Lazer, 2009). Δηλαδή, η πιθανότητα οι κοινωνικές σχέσεις ενός ατόμου αν μπορούν να οδηγήσουν σε επακόλουθες συμπεριφορές ταξιδιού και αντίστροφα.

Οι εξελίξεις στην έρευνα της οδηγικής συμπεριφοράς μπορούν ταυτόχρονα να προωθήσουν την έρευνα μεγάλων δεδομένων για την ανάλυση προτύπων ανθρώπινης κινητικότητας. Για παράδειγμα, έρευνα οδηγικής συμπεριφοράς έχει δείξει ότι οι στάσεις και οι προτιμήσεις ενός ατόμου παίζουν σημαντικό ρόλο στη λήψη αποφάσεων κατά τη διάρκεια του ταξιδιού (Garling, Gillholm, & Garling, 1998).

Οι στάσεις των ατόμων συχνά απαιτούνται από μια σειρά δηλώσεων ερευνητών, ενώ οι φυσικές μορφές που είναι διαθέσιμες από βάσεις μεγάλων δεδομένων (μέσα κοινωνικής δικτύωσης, π.χ. twitter) ενδέχεται να ανακαλύψουν νέες ιδέες σχετικά με τις στάσεις και τις αξίες και την ενίσχυση της μεθοδολογικής ανάπτυξης.

Επιπρόσθετα, ένα βήμα πέρα από την κατανόηση των συσχετισμών αλλά για την αποκάλυψη των μηχανισμών αιτιότητας θα ήταν χρήσιμο. Από την άποψη αυτή, τα ερευνητικά σχέδια και οι μέθοδοι όπου χρησιμοποιούνται σε ευρύ φάσμα στην έρευνα της οδηγικής συμπεριφοράς μπορούν να αναπτυχθούν για αποσαφήνιση της αιτιότητας (Bollen, 1989).

2.2 ΕΠΙΚΙΝΔΥΝΟΤΗΤΑ ΚΑΙ ΟΔΙΚΗ ΑΣΦΑΛΕΙΑ

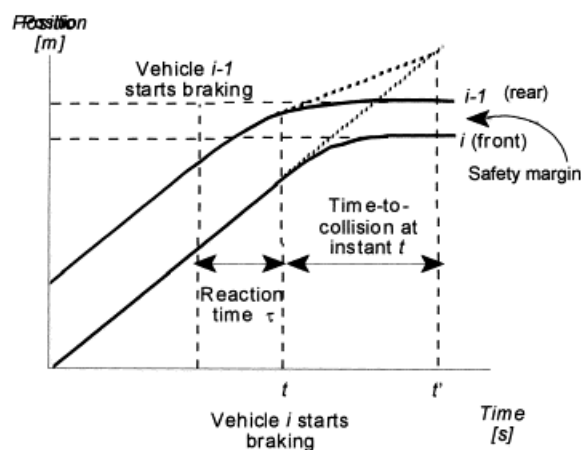
Η επικινδυνότητα κατά τη διάρκεια της οδήγησης αποτελεί ποσοτικοποίηση του κινδύνου πρόκλησης ατυχήματος. Πιο συγκεκριμένα, ορίζεται ως το γινόμενο επίπτωσης επί τον κίνδυνο και εξαρτάται άμεσα τόσο από την σοβαρότητα του επικείμενου συμβάντος (ατύχημα) όσο και από το πόσο πιθανό είναι αυτό (Assum, 1997).

Για την ποσοτικοποίηση λοιπόν, του όρου της επικινδυνότητας εισάγονται δείκτες επικινδυνότητας που παρουσιάζονται παρακάτω.

2.2.1 Δείκτες Επικινδυνότητας

Ο χρόνος σύγκρουσης (TTC – time to collision) θεωρείται ως ένας από τους πιο ευρέως διαδεδομένους δείκτες οδικής ασφάλειας και αποτελεί μια μέτρηση που εμφανίζει τη πιθανότητα εμπλοκής σε οδικό ατύχημα (Kiefer, Flannagan, & Jerome, 2006).

Ορίζεται ως ο χρόνος που απαιτείται για να συγκρουστούν δύο οχήματα εάν διατηρήσουν την τρέχουσα ταχύτητα και κατεύθυνση τους (Διάγραμμα 2. 1).



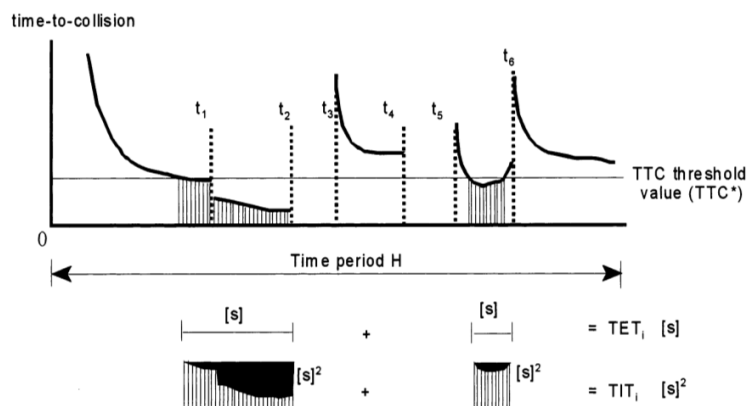
Διάγραμμα 2. 1: Ορισμός "χρόνου για κρούση" (Time to Collision – TTC)

Σημαντικός θεωρείται ο τρόπος αντίληψης του παραπάνω δείκτη από τον εκάστοτε χρήστη. Μέσω της κατανόησης του τρόπου αντίληψης του δείκτη TTC, διαμορφώνεται ο τρόπος οδήγησης κάθε χρήστη καθώς και το αν αυτός είναι επικίνδυνος ή όχι. Έπειτα από μια σειρά πειραμάτων με προσομοιωτή οδήγησης, κάθε οδηγός λαμβάνει υπόψιν του τρεις παράγοντες κατά την ερμηνεία του οδικού περιβάλλοντος (Hoffmann & Mortimer, 1994):

- Απόσταση από το αντικείμενο.
- Σχετική ταχύτητα των αντικειμένων.
- Χρόνο παρατήρησης.

Με βάση τον δείκτη TTC, διερευνήθηκαν δύο επιπλέον δείκτες ασφαλείας (Minderhoud & Bovy, 2001):

- **TET** (time exposed time-to-collision)
- **TIT** (time integrated time-to-collision)



Διάγραμμα 2. 2: Ορισμός και σχηματική απεικόνιση των δεικτών TET και TIT.
(Πηγή: (Minderhoud & Bovy, 2001))

Ο δείκτης TET δείχνει πόσο συνολικά διάστημα ένας οδηγός είναι εκτεθειμένος σε τιμές του TTC κάτω του ορίου επικινδυνότητας και συνεπώς εξαρτάται από το επιλεγθέν όριο για το TTC. Συνεπώς, όσο μεγαλύτερο το TET τόσο επικινδυνότερη η οδήγηση του χρήστη (Minderhoud & Bovy, 2001).

Το πρόβλημα ωστόσο του TET είναι ότι δεν επισημαίνει τη σοβαρότητα του εκάστοτε περιστατικού, παρά μόνο το εάν είναι ή όχι κάτω από το όριο ασφαλείας. Συμπληρωματικά λοιπόν εισήχθη ο δείκτης TIT, ο οποίος αντικατοπτρίζει το συνολικό εμβαδόν του διαγράμματος TTC – χρόνου ($TTC = f(t)$) το οποίο βρίσκεται κάτω από το επιλεγθέν όριο επικινδυνότητας.

2.2.2 Παράγοντες Επικινδυνότητας

Οι παράγοντες της επικινδυνότητας, σύμφωνα με τον Quimby (1981), αφορούν ουσιαστικά στα χαρακτηριστικά εκείνα των χρηστών της οδού τα οποία συντελούν στη διαμόρφωση συγκεκριμένης συμπεριφοράς από τους οδηγούς, η οποία μπορεί να είναι είτε επικίνδυνη (βάσει δεικτών επικινδυνότητας) είτε μη-επικίνδυνη. Βάσει έρευνας η οποία διεξήχθη από τους (Dingus et al. (2006), διαπιστώθηκε ότι παράγοντες επικινδυνότητας είναι μεταξύ άλλων:

- Το φύλο του οδηγού (άνδρας, γυναίκα)
- Η ηλικία του οδηγού
- Οι ταχύτητες της λοιπής κυκλοφορίας
- Η επιθετικότητα της οδήγησης του οδηγού
- Η οδήγηση με ταχύτητα μεγαλύτερη του ορίου ταχύτητας
- Ενασχόληση με το κινητό τηλέφωνο κατά τη διάρκεια της οδήγησης.

2.2.3 Είδη Ερευνών

Το θέμα της επικινδυνότητας και της ποσοτικοποίησης αυτής απασχολεί καθημερινά πολλούς μελετητές όσον αφορά έρευνες πάνω σε συγκοινωνιακά ζητήματα. Οι έρευνες αυτές μπορούν να διαχωριστούν βάσει του τρόπου προσέγγισης του προβλήματος και της συλλογής δεδομένων σε αναλυτικές, μετρήσεων πεδίου και πειραματικές μέσω προσομοίωσης σε μικροσκοπικό και μακροσκοπικό επίπεδο.

Οι έρευνες που σχετίζονται με την οδική ασφάλεια γίνονται κυρίως μέσω ερωτηματολογίων, μετρήσεων σε πραγματικό χρόνο με τη βοήθεια καταγραφικού εξοπλισμού ή με χρήση βάσεων δεδομένων ατυχημάτων. Πιο αναλυτικά, οι κατηγορίες ερευνών είναι οι εξής:

- *Αναλυτικές έρευνες*

Οι αναλυτικές έρευνες βασίζονται σε αναλυτικές εξισώσεις κίνησης προκειμένου να υπολογιστούν οι συντελεστές ασφαλείας και τα όριά τους (πχ. απαραίτητες αποστάσεις ορατότητας).

- *Έρευνες με Μετρήσεις Πεδίου*

Στην έρευνα των (Lauer and Forbes, 1930) έκαναν έναν “μικροσκοπικό αυτοκινητόδρομο” ως προσομοιωτή οδήγησης. Αυτός ο προσομοιωτής έγινε από μία ζώνη από καμβά, περίπου 20 πόδια μακριά, που επιτρέπει στον οδηγό της ίδιας κατεύθυνσης- και των αντίθετα κινουμένων μικροσκοπικών αυτοκινήτων-τον οπτικό έλεγχο του οδοστρώματος. Η έννοια αυτής της μικρογραφίας ήταν να μπορεί ο οδηγός να οδηγεί, να επιταχύνει και να επιβραδύνει για να προσπεράσει και να αποφύγει την κυκλοφορία διπλής κατεύθυνσης των μικροσκοπικών αυτοκινήτων. Αυτός ο προσομοιωτής βοήθησε στο να υπολογιστεί η ταχύτητα των αυτοκινήτων που είχαν προσπεραστεί και επίσης ο χρόνος προσπέρασης (Forbes, 1932). Οι Matson και Forbes (1938) ήταν πρωτοπόροι στο ότι αυτοί πρώτοι χρησιμοποίησαν κινούμενο όχημα παρατήρησης για να μελετήσουν την συμπεριφορά του οδηγού που προσπερνά. Το όχημα παρατήρησης οδηγείται κατά μήκος της κυκλοφοριακής ροής με προκαθορισμένη ταχύτητα και καταγράφει τη συμπεριφορά των οδηγών που το προσπερνά. Εκτός από το όχημα παρατήρησης, χρησιμοποίησαν επίσης μια οπτική μέθοδο καταγραφής της συμπεριφοράς των αυτοκινητιστών ενώ εκτελούσαν την προσπέραση. Αυτή η μέθοδος βοήθησε στην ανάλυση των στοιχείων στο εργαστήριο. Ανέλυσαν περίπου 709 προσπεράσεις. Μια νέα μέθοδος εφαρμόστηκε στην έρευνα για τη συμπεριφορά κατά την προσπέραση, χρησιμοποιώντας πρόσθετα όργανα παρατήρησης στο εξεταζόμενο όχημα.

- *Έρευνες με Προσομοιωτή Οδήγησης*

Οι προσομοιωτές οδήγησης βασίζονται σε μικροσκοπικά πρότυπα, καθώς η περιγραφή τους δεν εντοπίζεται μόνο στις ίδιες τις μονάδες του συστήματος, δηλαδή τα οχήματα, αλλά και στις αλληλεπιδράσεις αυτών, γεγονός που διευκολύνει σαφώς τη γρήγορη και εύκολη συγκέντρωση στοιχείων και μεταβλητών που αφορούν στην προσπέραση. Αρκετές έρευνες έχουν βασιστεί σε πειράματα από προσομοιωτές οδήγησης, επειδή δεν περιέχουν τον ρεαλιστικό κίνδυνο και η καταγραφή καταστάσεων που κρίνονται δύσκολες σε πραγματικό χρόνο είναι εφικτή. Για παράδειγμα, πιθανότητα εμπλοκής οχήματος σε ατύχημα.

2.2.4 Εργαλείο Επισήμανσης Επικίνδυνων Σημείων

Όπως παρουσιάζεται από τους (Weekley et al., 2016), το ESReT (European Safety Review Tool) είναι ένα εργαλείο υποστήριξης των οδικών τμημάτων τα οποία έχουν χαρακτηριστεί ως επικίνδυνα. Οι αλγόριθμοι οδικής ασφάλειας αποτελούν ένα από τα βασικά συστατικά του εργαλείου που αναπτύσσεται. Επιτρέπουν τα δεδομένα που συλλέγονται βάσει των χαρακτηριστικών της οδού να συγκρίνονται με τις ιδανικές τιμές σύμφωνα με την ασφάλεια. Το εργαλείο αυτό συμβάλλει στην επισήμανση τμημάτων υψηλού κινδύνου, παρέχοντας με τον τρόπο αυτό βοήθεια στους ειδικούς που μελετούν θέματα οδικής ασφάλειας.

Οι αλγόριθμοι χρησιμοποιούν μια σειρά χαρακτηριστικών και συνθηκών για να προσδιορίσουν τη σοβαρότητα μιας δεδομένης τοποθεσίας:

- Ανεπεξέργαστα χαρακτηριστικά, στοιχεία που συλλέγονται από έρευνες με κάμερες ή με αυτόματη εικόνα.
- Υπολογισμένα χαρακτηριστικά, τα οποία χρησιμοποιούνται στις ιδανικές τιμές ασφαλείας του εργαλείου.
- Διακόπτες ασφαλείας
- Αντιμετώπιση
- Πληροφόρηση

Το εργαλείο είναι προσαρμοσμένο ώστε να παρέχει μια σειρά βασικών λειτουργιών στους χρήστες:

- Επιτρέπει στους χρήστες να εγγράφονται σε διαφορετικά επίπεδα πρόσβασης.
- Επιτρέπει στους χρήστες να ελέγχουν, να επεξεργάζονται και να προσθέτουν νέα πρότυπα σχεδίασης.
- Είναι ικανό να συλλέγει δεδομένα από διαφορετικές πηγές.
- Επιτρέπει στους χρήστες να ελέγχουν τα δεδομένα μέσω εικόνων ή χαρτογράφησης.
- Διευκολύνει τη σύγκριση των δεδομένων των ερευνών με τους κανόνες σχεδιασμού και ασφαλών συστημάτων προκειμένου να εντοπιστούν, να επισημανθούν και να παρουσιασθούν τυχόν ελλείμματα.
- Παρέχει πληροφορίες σχετικά με τις πιθανές επιλογές αντιμετώπισης των ελλειμμάτων που εντοπίστηκαν.

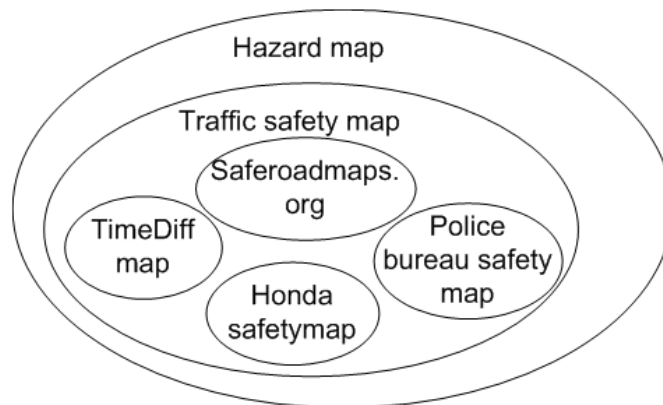
Στην Εικόνα 2. 1 φαίνεται το περιβάλλον του εργαλείου. Πιο συγκεκριμένα, η διαδρομή που ακολουθήθηκε από το χρήστη και στο γράφημα απεικονίζεται η διακύμανση της ταχύτητας του οδηγού.



Εικόνα 2. 1: Περιβάλλον εφαρμογής (Πηγή:(Weekley et al., 2016)).

2.2.5 Χάρτες Οδικής Ασφάλειας

Οι χάρτες οδικής ασφάλειας είναι ένα είδος ανάλυσης επικινδυνότητας του οδικού δικτύου. Σήμερα, οι περισσότεροι χάρτες οδικής ασφάλειας αποτελούν απεικόνιση ατυχημάτων από στοιχεία της αστυνομίας και είναι προσβάσιμοι μέσω διαδικτύου για τους πολίτες. Δείχνουν στοιχεία όπως την τοποθεσία του ατυχήματος, τον αριθμό των εμπλεκόμενων οχημάτων ή τον τύπο της σύγκρουσης. Η διαδικασία αυτή εκτός από χρονοβόρα, κοστίζει αρκετά. Έτσι, με την εμφάνιση των έξυπνων κινητών τηλεφώνων η κατάσταση μπορεί να απλοποιηθεί. Στην Εικόνα 2. 2 που ακολουθεί, φαίνεται πως ένας χάρτης οδικής ασφάλειας είναι ένας τύπος χάρτη επικινδυνότητας, ο οποίος περιλαμβάνει πληροφορίες για το περιβάλλον με στόχο την πληροφόρηση των οδηγών.



Εικόνα 2. 2: Χάρτης Επικινδυνότητα και Χάρτης Οδικής Ασφάλειας (Πηγή:(Dang, Kubo, Sato, & Namatame, 2015)).

Οι προϋποθέσεις για έναν χάρτη που κατασκευάζεται από δεδομένα έξυπνων κινητών τηλεφώνων είναι οι εξής:

1. *Ακρίβεια.* Η ακρίβεια εξαρτάται από την αναγνώριση συχνών συμβάντων σε συγκεκριμένες τοποθεσίες. Οι θέσεις και η συχνότητα των ατυχημάτων πρέπει να περιγράφονται με ακρίβεια.
2. *Αυτόματη δημιουργία.* Επειδή η δημιουργία τέτοιων χαρτών απαιτεί πολύ χρόνο και προσπάθεια, είναι χρήσιμο η διαδικασία να γίνεται αυτόματα.
3. *Στοιχεία Ιδιωτικού Απορρήτου.* Δεν πρέπει να διαφεύγουν προσωπικά δεδομένα των χρηστών.
4. *Συμμετοχή της Κοινότητας.* Με τη συμμετοχή πολλών χρηστών ο χάρτης γίνεται ακριβής και καλύπτει μεγαλύτερες εκτάσεις.
5. *Απεικόνιση των διαφορών.* Οι διαφορές με τις οποίες οι χρήστες αντιλαμβάνονται το περιβάλλον πρέπει να απεικονίζονται στο χάρτη.

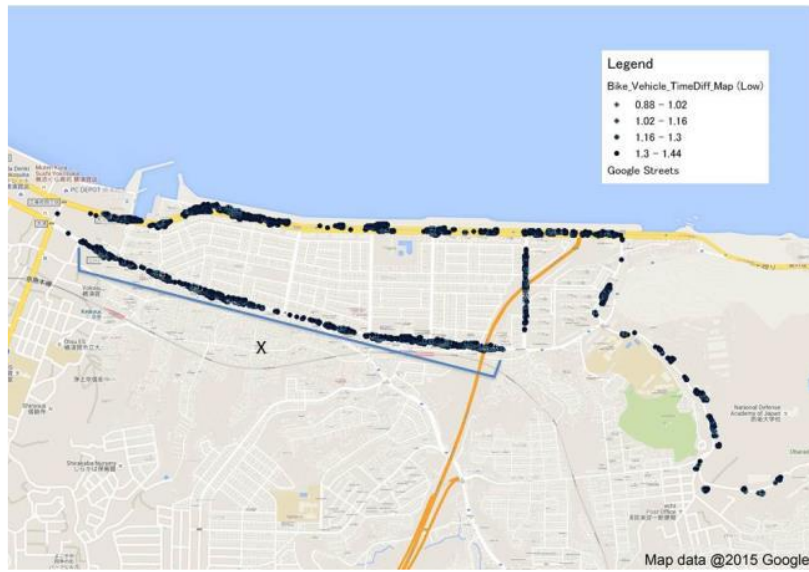
6. *Απεικόνιση των συμμετεχόντων.* Είναι χρήσιμο να γίνεται διάκριση πολυσύχναστων και μη πολυσύχναστων δρόμων πάνω στο χάρτη.
7. *Χάρτης Ασφάλειας για όλους.* Ο χάρτης που θα δημιουργηθεί δεν πρέπει να απευθύνεται μόνο σε οδηγούς οχημάτων.

Οι (Dang et al., 2015) έχουν εισάγει μια μέθοδο για τη δημιουργία χαρτών ασφαλείας από οδικά συμβάντα που μπορεί να χρησιμοποιηθεί για τον τρόπο με τον οποίο ο οδηγός αντιλαμβάνεται το περιβάλλον. Κατασκευάζουν ένα μοντέλο για το σκοπό αυτό.

Σε έρευνα που πραγματοποιήθηκε στην Ιαπωνία, για οκτώ χιλιόμετρα δρόμου σε διάστημα τριάντα ημερών, προκύπτουν οι παρακάτω χάρτες ασφαλείας. Στον χάρτη της Εικόνας 2. 3 απεικονίζονται τα σημεία υψηλού κινδύνου, ενώ στο χάρτη της Εικόνας 2. 4 τα σημεία χαμηλού κινδύνου.



Εικόνα 2. 3: Χάρτης Οδικής Ασφάλειας-Περιοχή Υψηλού Κινδύνου (Πηγή:(Dang et al., 2015)).



Εικόνα 2. 4:Χάρτης Οδικής Ασφάλειας-Περιοχή Χαμηλού Κινδύνου (Πηγή:(Dang et al., 2015)).

2.3 ΟΜΑΔΟΠΟΙΗΣΗ ΣΥΜΠΕΡΙΦΟΡΑΣ

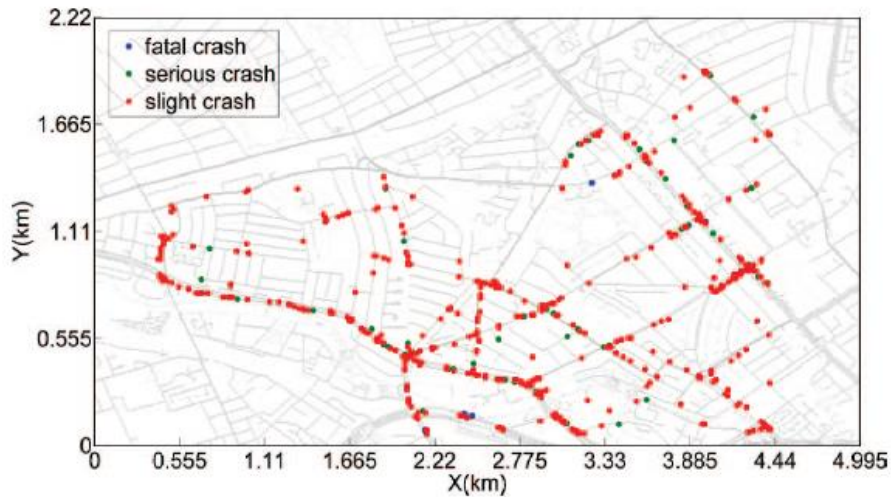
Στις περισσότερες περιπτώσεις τα οδηγικά ατυχήματα σχηματίζουν ομάδες στον χώρο, οι οποίες ονομάζονται clusters ή hotspots. Η ομαδοποίηση αυτή περιλαμβάνει τέσσερις φάσεις: αναγνώριση, ταξινόμηση, χαρακτηρισμό και επεξεργασία και αποτελεί διαδικασία κλειδί για συγκοινωνιακά ζητήματα.

Αρχικά, το πρόβλημα της αναγνώρισης περιλαμβάνει τον εντοπισμό του είδους της ανάλυσης που πρόκειται να χρησιμοποιηθεί, δηλαδή στην ουσία η επιλογή του καταλληλότερου αλγορίθμου ομαδοποίησης για το εκάστοτε πρόβλημα.

Η μια κατηγορία αλγορίθμων αναφέρεται σε δημιουργία ομάδων συγκεκριμένων σχημάτων (π.χ. κύκλος ή τετράγωνο) και η άλλη δημιουργεί αυθαίρετα σχήματα στο χώρο. Ωστόσο, και οι δύο κατηγορίες αλγορίθμων είναι ευαίσθητες όσον αφορά την επιλογή των κατάλληλων παραμέτρων.

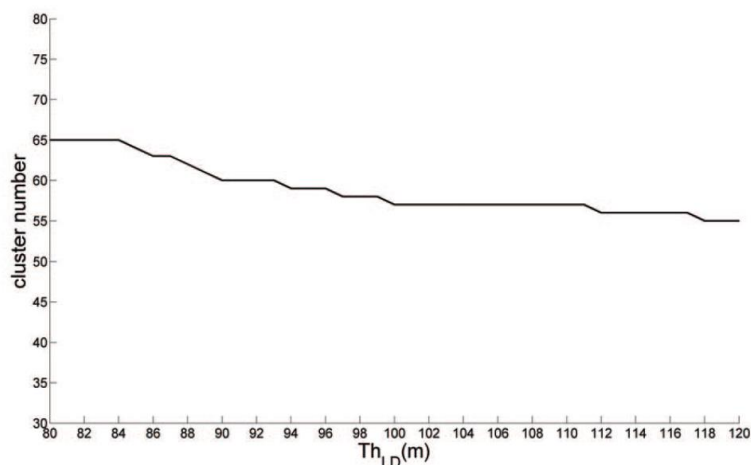
Για να ξεπεραστεί το πρόβλημα αυτό, οι (Zou & Zhu, 2011) προτείνουν μια μέθοδο που ονομάζεται Natural Nearest Neighbor (3N), μέσω της οποίας γειτονικές ομάδες γίνονται αποδεκτές μεταξύ τους σύμφωνα με τις σχέσεις τους. Η μέθοδος αυτή, δεν απαιτεί την αρχική εισαγωγή παραμέτρων από το χρήστη. Αντί της ευκλείδειας απόστασης συσσωρεύει χωρικά σημεία ανάλογα με την απόσταση του δικτύου και είναι ικανή να δημιουργήσει ομάδες αυθαίρετων σχημάτων. Το μόνο μειονέκτημα της μεθόδου είναι πως εξ αιτίας του γεγονότος ότι μπορεί να καλύψει μια μεγάλη περιοχή ανάλυσης, είναι πιθανό να οδηγήσει σε δύσκολες θέσεις των hotspot σημείων, των ομάδων δηλαδή.

Πραγματοποιήθηκε έρευνα από τους (Han et al., 2015) για ένα δείγμα οδηγικών συγκρούσεων στο Ηνωμένο Βασίλειο για τα έτη από το 2008 έως το 2013, για εφαρμογή της παραπάνω μεθόδου ανίχνευσης ομάδων. Τα δεδομένα αφορούν συντεταγμένες σημείων όπου πραγματοποιήθηκε μια σύγκρουση και απεικονίζονται στο χάρτη της Εικόνας 2. 6.



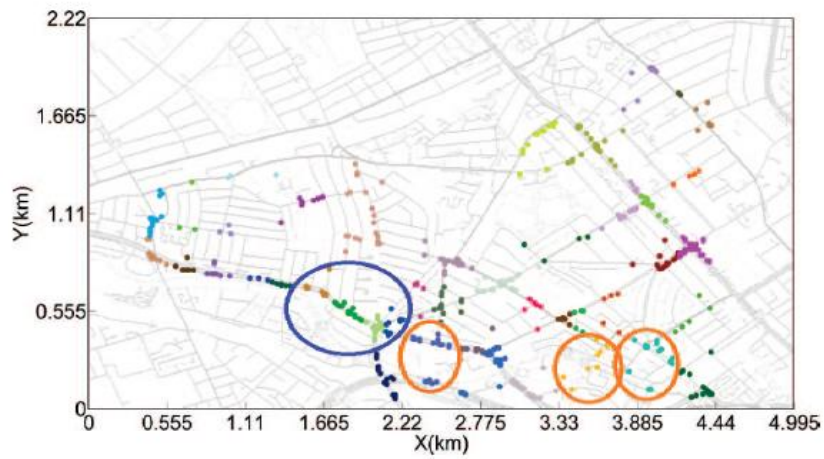
Εικόνα 2. 5: Κατανομή συγκρούσεων.

Όσον αφορά το πρώτο πείραμα, επιλέγεται μια ακτίνα επιρροής ατυχημάτων και κατασκευάζεται ένα διάγραμμα (Διάγραμμα 2. 3) συσχέτισης της μεταβλητής αυτής με τον αριθμό των ομάδων που δημιουργούνται κάθε φορά. Όπως παρατηρείται δεν υπάρχει ουσιαστική συσχέτιση μεταξύ των δύο μεταβλητών.

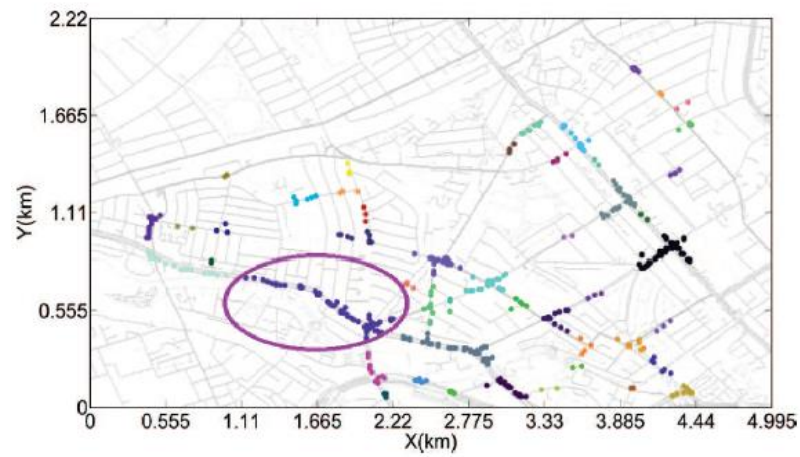


Διάγραμμα 2. 3: Συσχέτιση μεταξύ πλήθους ομάδων και ακτίνας.

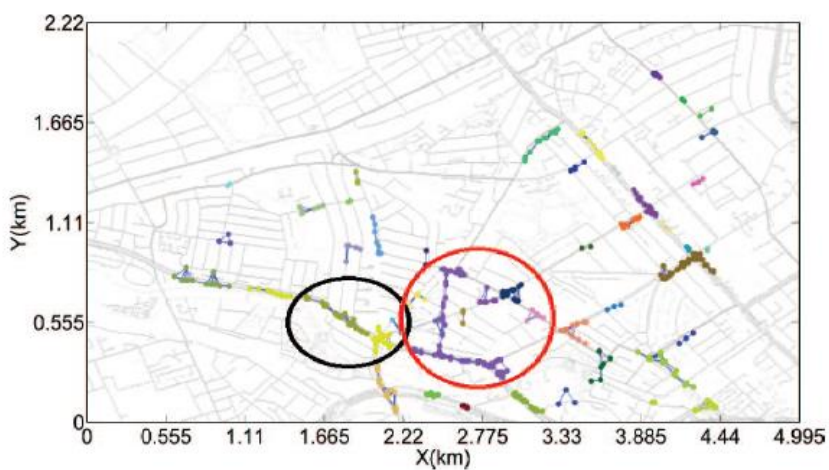
Στο δεύτερο πείραμα, γίνεται εφαρμογή αλγορίθμων ομαδοποίησης για τα ίδια δεδομένα. Τα αποτελέσματα από τους αλγορίθμους αυτούς παρουσιάζονται στις παρακάτω εικόνες. Κάθε διαφορετικό χρώμα αντιπροσωπεύει διαφορετική ομάδα.



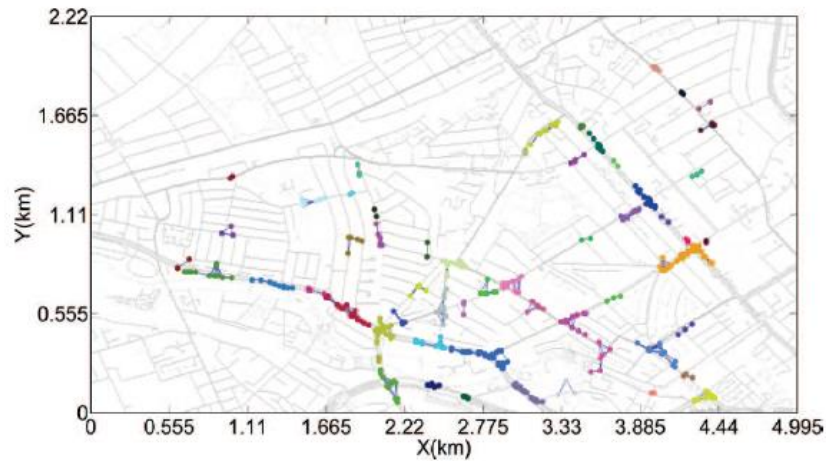
Εικόνα 2. 6: Αλγόριθμος K-means.



Εικόνα 2. 7: Αλγόριθμος DBSCAN.



Εικόνα 2. 8: Αλγόριθμος CB3N.



Εικόνα 2. 9: Αλγόριθμος DHT3N.

Η προτεινόμενη μέθοδος, DHT3N, χωρίζει την κόκκινη έλλειψη της μεθόδου CB3N σε τέσσερις διαφορετικές ομάδες. Όλες οι ομάδες απεικονίζουν το χαρακτήρα της οδού, το οποίο διακρίνει τη διασταύρωση ή όχι.

Αντίστοιχη έρευνα έχει γίνει και από τους (Ranjith, Athanesious, & Vaidehi, 2016), για ανίχνευση ανωμαλιών στο οδικό δίκτυο με εφαρμογή του αλγορίθμου DBSCAN. Μέσω της διαδικασίας αυτής, ανιχνεύονται σημεία ανωμαλιών, τα οποία δεν έχουν καμία σχέση με τα υπόλοιπα σημεία του δικτύου και δεν εντάσσονται σε κάποια από τις ομάδες. Ο αλγόριθμος βρίσκει σημεία υψηλού και χαμηλού κινδύνου. Τα αποτελέσματα της έρευνας επιτυγχάνονται με ποσοστό ακρίβειας ίσο με 68.70%.

2.4 ΣΥΜΠΕΡΑΣΜΑΤΑ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Από την ανάλυση της βιβλιογραφίας γίνεται αντιληπτή η σημασία των δεδομένων από αισθητήρες έξυπνων κινητών τηλεφώνων για την ανάλυση της οδηγικής συμπεριφοράς. Τα δεδομένα από τα κινητά τηλέφωνα φέρουν αξιόπιστα αποτελέσματα και συμβάλλουν στην αντιμετώπιση συγκοινωνιακών ζητημάτων.

Η αναγνώριση της οδηγικής συμπεριφοράς αποτελεί ζήτημα που απασχολεί έντονα πολλούς κλάδους της επιστήμης. Στην βιβλιογραφία υπάρχουν αρκετές αναφορές για τους παράγοντες που επηρεάζουν την οδηγική συμπεριφορά αλλά και τη λήψη αποφάσεων κατά τη διάρκεια της οδήγησης από τους οδηγούς.

Όσον αφορά την επικινδυνότητα κατά τη διάρκεια της οδήγησης, οι περισσότερες έρευνες αφορούν δείκτες επικινδυνότητας και παράγοντες που συμβάλλουν στην επικίνδυνη οδήγηση. Οι δείκτες αυτοί συνήθως εκφράζονται σε όρους ατυχημάτων. Ωστόσο, η επικίνδυνη οδήγηση δεν οδηγεί μόνο σε ατύχημα. Για το λόγο αυτό, κρίνεται αναγκαία η ανάλυση της οδηγικής συμπεριφοράς μέσα από πληθοποριστικά δεδομένα προκειμένου να βρεθούν τμήματα του δικτύου που παρουσιάζεται γενικότερη επικίνδυνη συμπεριφορά.

Επιπλέον, η βιβλιογραφία κατέστη επαρκής όσον αφορά την απεικόνιση των σημείων όπου γίνεται μια σύγκρουση. Ωστόσο, τα επικίνδυνα σημεία σε ένα οδικό δίκτυο δεν αποτελούνται μόνο από τα ατυχήματα. Ένα επικίνδυνο σημείο μπορεί να είναι ένα σημείο του δικτύου το οποίο οι οδηγοί κάνουν συχνά λάθος. Έτσι, είναι χρήσιμο να πραγματοποιηθεί απεικόνιση των ακραίων οδηγικών συμβάντων. Με τον τρόπο αυτό ο χρήστης είναι ενήμερος για την οδηγική του συμπεριφορά και για τα λάθη τα οποία πραγματοποιεί κατά τη διάρκεια της οδήγησης, προκειμένου να μπορεί να λαμβάνει εγκαίρως τις σωστές αποφάσεις για αποφυγή παρόμοιας μελλοντικής ακραίας συμπεριφοράς. Επιπρόσθετα, μέσω της διαδικασίας αυτής το σύστημα διαχείρισης της κυκλοφορίας και οδικής ασφάλειας είναι σε θέση να γνωρίζει τα επικίνδυνα σημεία του οδικού δικτύου, με σκοπό να προβεί σε ενέργειες βελτίωσης του.

3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Το παρόν κεφάλαιο αποσκοπεί σε μια αναλυτική περιγραφή των μεθοδολογιών που χρησιμοποιήθηκαν στην προσπάθεια σύνθεσης της έρευνας. Με τον τρόπο αυτό, ο αναγνώστης είναι σε θέση να παρακολουθήσει την εξέλιξη της παρούσας έρευνας, χωρίς να ανατρέξει σε άλλες πηγές προκειμένου να κατανοήσει την μεθοδολογία που χρησιμοποιήθηκε. Αρχικά, παρουσιάζεται το διάγραμμα ροής των εργασιών της έρευνας και σε επόμενο στάδιο γίνεται περιγραφή των αλγορίθμων που χρησιμοποιήθηκαν για την αντιμετώπιση του προβλήματος.

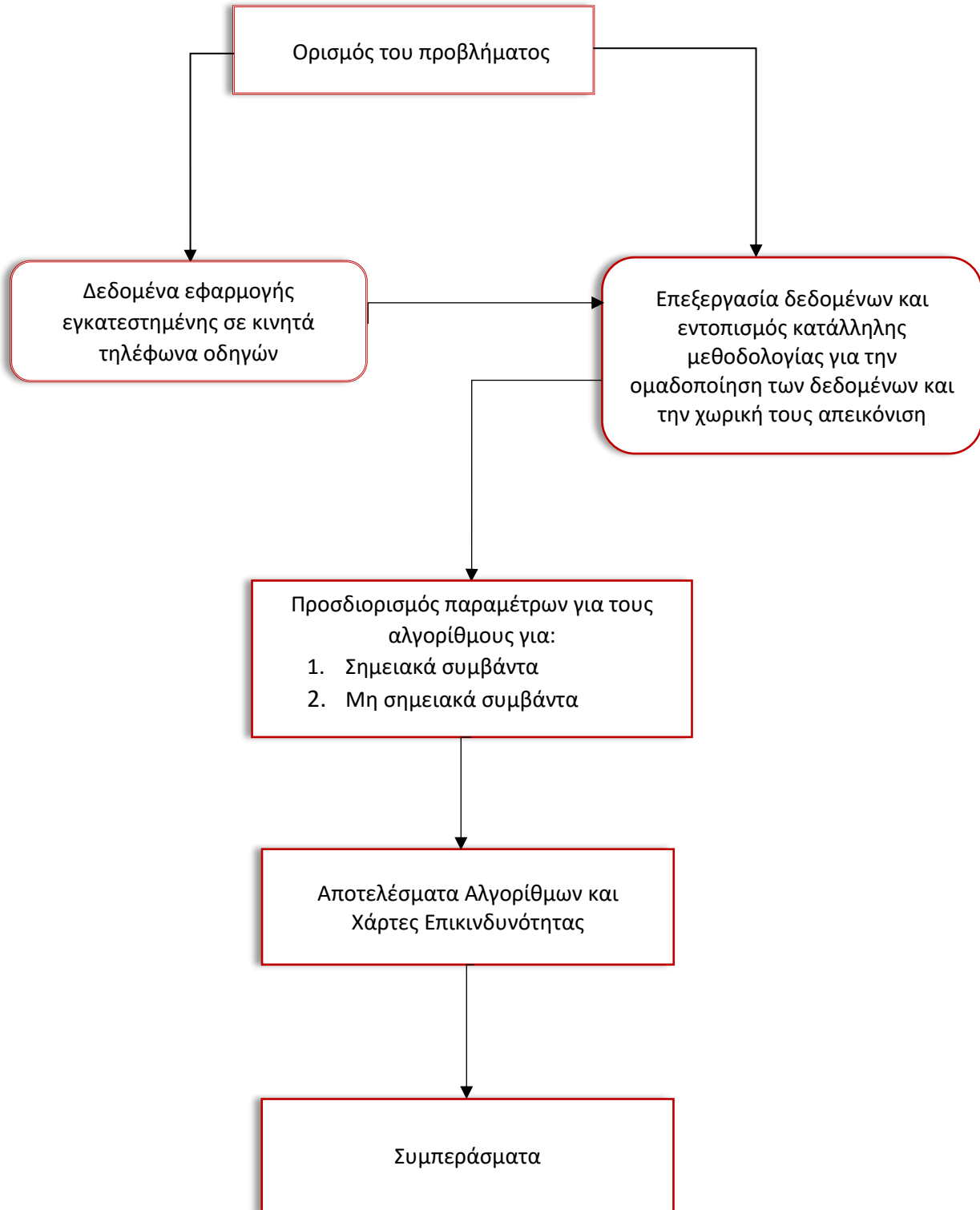
3.1 ΠΡΟΣΕΓΓΙΣΗ ΚΑΙ ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ ΕΡΓΑΣΙΩΝ

Στην παρούσα εργασία επιδιώκεται η ομαδοποίηση της συμπεριφοράς των οδηγών κατά τη διάρκεια ενός ταξιδιού με στόχο την έγκαιρη προειδοποίηση (πριν την εκκίνηση ενός νέου ταξιδιού) των χρηστών για βελτίωση της οδηγικής τους συμπεριφοράς. Τα οδικά χαρακτηριστικά στα οποία βασίστηκε η έρευνα είναι τα ακόλουθα:

- Η συχνότητα απότομων επιταχύνσεων και επιβραδύνσεων (επιθετική οδήγηση)
- Το ποσοστό χρόνου οδήγησης στο οποίο βρίσκεται άνω του ορίου ταχύτητας.

Τα παραπάνω στοιχεία αντλούνται από βάση δεδομένων της εφαρμογής κινητού τηλεφώνου OSeven. Έπειτα, προσδιορίζονται τα απαραίτητα στοιχεία για το «στήσιμο» του προβλήματος. Όσον αφορά σημειακά συμβάντα στη συμπεριφορά του οδηγού (απότομη επιτάχυνση και επιβράδυνση) έγινε χρήση του αλγορίθμου DBSCAN για τον εντοπισμό και την ένταξη των δεδομένων σε συστάδες. Για τα γεγονότα που αφορούν την υπέρβαση του ορίου ταχύτητας κατά τη διάρκεια της οδήγησης δεν είναι δυνατή η χρήση του ίδιου αλγορίθμου καθώς πρόκειται για χωροχρονικά πλέον δεδομένα. Ο αλγόριθμος DBSCAN δεν λαμβάνει υπόψιν του τον παράγοντα του χρόνου στην ανάλυση των δεδομένων. Για τον λόγο αυτό, χρησιμοποιείται ο αλγόριθμος Apriori και ένας αλγόριθμος εξόρυξης προτύπων ακολουθιών (frequent sequence mining), μέσω του οποίου εντοπίζονται μοτίβα (patterns) από σημεία που βρίσκονται και παρουσιάζουν ένα συμβάν στη σειρά.

Στο Διάγραμμα 3. 1 περιγράφεται εικονικά η ροή των εργασιών στο πλαίσιο της παρούσας εργασίας.



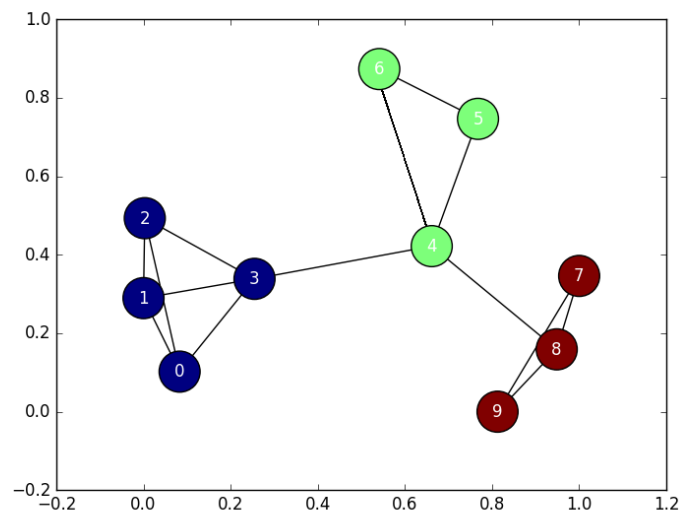
Διάγραμμα 3. 1: Διάγραμμα ροής εργασιών διπλωματικής εργασίας.

3.2 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

3.2.1 Ομαδοποίηση

Μια ομάδα ή συστάδα είναι μια συλλογή από δεδομένα τα οποία εμφανίζουν παρόμοια συμπεριφορά μεταξύ τους και διαφέρουν από δεδομένα που βρίσκονται σε κάποια άλλη ομάδα. Έτσι, η ομάδα αυτών των δεδομένων μπορεί να συμπεριφερθεί και να επεξεργαστεί συλλογικά. Η διαδικασία έρευνας των ομάδων σε μια βάση δεδομένων καλείται ομαδοποίηση (clustering).

Στην Εικόνα 3. 1, απεικονίζονται τρεις διαφορετικές ομάδες. Όπως είναι εμφανές, η ομάδα με το μπλε χρώμα αποτελείται από τα στοιχεία 0,1,2,3, η ομάδα με το πράσινο χρώμα αποτελείται από τα στοιχεία 4,5,6 και η ομάδα με το κόκκινο χρώμα αποτελείται από τα στοιχεία 7,8,9. Συνεπώς, κάθε μια από τις ομάδες αυτές περιλαμβάνει κοντινά σημεία, τα οποία πιθανόν να παρουσιάζουν όμοια συμπεριφορά στο χώρο.



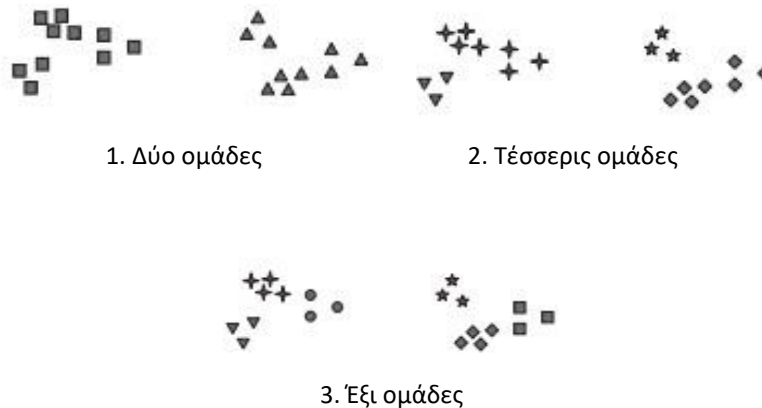
Εικόνα 3. 1: Γραφική απεικόνιση τριών ομάδων.

Η ομαδοποίηση ως διαδικασία μπορεί ακόμα να χρησιμοποιηθεί για την ανίχνευση απομακρυσμένων σημείων (outliers, σημεία τα οποία είναι αρκετά μακριά από κάποια ομάδα), τα οποία εμφανίζουν ιδιαίτερο ενδιαφέρον σε σχέση ακόμα και με τις κοινές περιπτώσεις αντικειμένων.

Η ακριβής σημασία της έννοιας “cluster” δεν είναι εύκολο να δοθεί. Προκειμένου να γίνει πιο κατανοητή η έννοια αυτή παρουσιάζονται τα παρακάτω σχήματα. Στο πρώτο φαίνονται τα είκοσι αρχικά σημεία, στα οποία θα γίνει ο διαχωρισμός σε διαφορετικές ομάδες και τα τρία επόμενα δείχνουν από έναν διαφορετικό τρόπο για την ομαδοποίηση αυτή.



Εικόνα 3. 2: Σύνολο σημείων για ομαδοποίηση.



Εικόνα 3. 3: Τρόποι διαχωρισμού σημείων σε ομάδες.

Στις εικόνες 1,2 και 3 γίνεται διάκριση των δεδομένων σε δύο, τέσσερα και έξι μέρη αντίστοιχα. Μέσα από αυτή την απεικόνιση γίνεται αντιληπτή η δυσκολία στην αντίληψη της έννοιας των cluster. Ο προσδιορισμός της εξαρτάται από τη φύση των δεδομένων που πρόκειται να αναλυθούν, καθώς και από τα αποτελέσματα που ο εκάστοτε ερευνητής επιθυμεί να επεξεργαστεί.

Στο πλαίσιο της παρούσας εργασίας, ομάδα θεωρείται ένα σύνολο από συντεταγμένες, στην οποία παρατηρείται ομοιότητα στη συμπεριφορά των οδηγών στο οδικό δίκτυο της Αθήνας.

Οι αλγόριθμοι που αφορούν την ομαδοποίηση δεδομένων μπορούν να κατηγοριοποιηθούν σε πέντε τύπους:

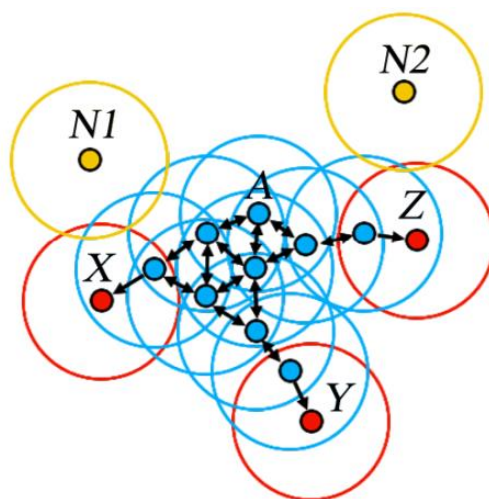
1. **Partitional - K-Means** . Ο αλγόριθμος εκτελεί συσταδοποίηση αντικειμένων στον n -διάστατο ευκλείδειο χώρο ενώ συνηθέστερα η ευκλείδεια νόρμα χρησιμοποιείται ως συνάρτηση απόστασης. Το μειονέκτημα της μεθόδου είναι πως πριν την εφαρμογή του απαιτεί τον καθορισμό το αριθμό των clusters να έχει καθοριστεί από το χρήστη. Είναι σημαντικό να αναφερθεί ότι ο k-Means είναι ευαίσθητος στο θόρυβο και την ύπαρξη ακραίων τιμών στα δεδομένα (outliers).
2. **Hierarchical - CURE BIRCH**. Αλγόριθμοι τέτοιου τύπου δημιουργούν ομάδες από γειτονικά clusters, τα οποία σχηματίζουν ένα δέντρο ιεραρχίας. Κάθε κόμβος από το δέντρο αντιπροσωπεύει ένα cluster από τη βάση δεδομένων.
3. **Grid- based - STING AND WaveCluster**. Βασίζονται σε δομές δικτύου πολλαπλών δεδομένων, στις οποίες εκτελούνται όλες οι απαραίτητες λειτουργίες για ομαδοποίηση.
4. **Model-based - COB-WEB**. Για κάθε cluster θεωρείται ένα μοντέλο και η βασική ιδέα είναι να βρεθεί η καλύτερη εφαρμογή του μοντέλου αυτού στα υπόλοιπα. Συχνά βασίζεται στην παραδοχή ότι τα δεδομένα παράγονται από κατανομές πιθανοτήτων.
5. **Density-based – DBSCAN, OPTICS, DENCLUE, CURD**. Βασίζονται στην ιδέα ότι τα δεδομένα τα οποία ορίζουν μια πυκνή περιοχή σχηματίζουν ένα cluster. Χρησιμοποιείται ένα ρυθμιζόμενο όριο που καθορίζει την εκάστοτε πυκνότητα των περιοχών. Αναζητά περιοχές υψηλής πυκνότητας με συγκεκριμένα χαρακτηριστικά γνωρίσματα.

3.2.2 Αλγόριθμος DBSCAN

Στα πλαίσια της παρούσας διπλωματικής εργασίας έγινε χρήση του αλγορίθμου DBSCAN για την επεξεργασία των σημειακών συμβάντων και την μετέπειτα αξιολόγηση των αποτελεσμάτων. Βασικά στοιχεία του αλγορίθμου παρουσιάζονται στο κεφάλαιο αυτό.

Πρόκειται να γίνει συσταδοποίηση για ένα σύνολο δεδομένων στο χώρο. Προκειμένου να πραγματοποιηθεί η συσταδοποίηση αυτή μέσω του αλγορίθμου DBSCAN, τα σημεία αυτά κατηγοριοποιούνται σε κεντρικά (core points), προσεγγίσιμα ή πυκνά-προσεγγίσιμα (density-reachable points) ή ακραία (outliers) με βάση τους παρακάτω κανόνες:

1. Ένα σημείο p είναι κεντρικό σημείο, αν τουλάχιστον $MinPts$ σημεία βρίσκονται σε απόσταση ϵ από αυτό και αυτά τα σημεία είναι άμεσα προσεγγίσιμα από το p (σημεία με γαλάζιο χρώμα). Κανένα σημείο δεν είναι προσεγγίσιμο από μη κεντρικό σημείο.
2. Ένα σημείο q είναι πυκνά-προσεγγίσιμο από ένα σημείο p , αν υπάρχει μονοπάτι p_1, \dots, p_n , με $p_1 = p$ και $p_n = q$, όπου κάθε $p_i + 1$ είναι άμεσα προσεγγίσιμο από το p_i , δηλαδή όλα τα σημεία του μονοπατιού πρέπει να είναι κεντρικά, με πιθανή εξαίρεση το q (σημεία X, Y και Z).
3. Σημεία, τα οποία δεν είναι προσεγγίσιμα από κανένα άλλο σημείο, είναι ακραία (σημεία $N1$ και $N2$).



Εικόνα 3. 4: Κατηγοριοποίηση σημείων.

Αν το σημείο p είναι κεντρικό σημείο, τότε σχηματίζει μια συστάδα μαζί με όλα τα σημεία (κεντρικά ή μη), τα οποία είναι προσεγγίσιμα από αυτό. Κάθε συστάδα περιέχει τουλάχιστον ένα κεντρικό σημείο. Η προσεγγισιμότητα δεν είναι συμμετρική σχέση, καθώς εξ ορισμού μπορεί κανένα σημείο να μην είναι προσεγγίσιμο από ένα μη κεντρικό σημείο, ανεξάρτητα από την απόσταση μεταξύ τους. Δηλαδή, ένα μη κεντρικό σημείο μπορεί να είναι προσεγγίσιμο, αλλά κανένα σημείο να μην μπορεί να προσεγγιστεί από αυτό. Συνεπώς, μια επιπλέον έννοια διασύνδεσης απαιτείται για τον ορισμό των συστάδων που δημιουργεί ο DBSCAN. Δυο σημεία p και q είναι πυκνά-συνδεδεμένα, αν υπάρχει σημείο s , τέτοιο ώστε τα p και q να είναι (πυκνά-προσεγγίσιμα) από το σημείο s . Η πυκνή-συνδεσιμότητα είναι συμμετρική σχέση.

Έτσι, μια συστάδα εμφανίζει τις δυο ακόλουθες ιδιότητες:

1. Όλα τα σημεία μιας συστάδας είναι αμοιβαία πυκνά-συνδεδεμένα μεταξύ τους.
2. Αν ένα σημείο είναι πυκνά-προσεγγίσιμο από κάποιο σημείο της συστάδας, τότε είναι και αυτό, επίσης, σημείο της συστάδας.

Ο αλγόριθμος DBSCAN προκειμένου να βρει μια ομάδα χρησιμοποιεί δυο παραμέτρους, το ϵ και το MinPts, δηλαδή την ακτίνα της κάθε συστάδας (η μεγαλύτερη απόσταση που θα απέχουν τα σημεία της ίδιας συστάδας) και τον ελάχιστο αριθμό σημείων που απαιτούνται για τη δημιουργία μιας πυκνής περιοχής αντίστοιχα. Ξεκινάει από ένα τυχαίο σημείο, το οποίο δεν έχει επισκεφθεί. Για το επιλεγμένο σημείο ανακτώνται τα σημεία στην ϵ -γειτονιά, δηλαδή σημεία που απέχουν το πολύ απόσταση ϵ από το επιλεγμένο σημείο. Αν υπάρχει επαρκής αριθμός σημείων, δηλαδή μεγαλύτερος του MinPts, ο αλγόριθμος δημιουργεί μια συστάδα. Σε αντίθετη περίπτωση, το σημείο επισημαίνεται προσωρινά ως θόρυβος(outlier). Μαρκάζονται προσωρινά, διότι σημεία που έχουν μαρκιαστεί ως θόρυβος, μπορεί στην πορεία να βρεθούν σε κάποια ϵ -γειτονιά άλλου επιλεγμένου σημείου και έτσι τελικά να γίνουν μέρος κάποιας συστάδας. Αν ένα σημείο αποτελεί πυκνό τμήμα μιας συστάδας, τότε σίγουρα η ϵ -γειτονιά του είναι υποσύνολο της συστάδας αυτής. Έτσι, όλα τα σημεία μέσα στην ϵ -γειτονιά προστίθενται στη συστάδα, καθώς επίσης και τα σημεία στην ϵ -γειτονιά καθενός από αυτά τα σημεία.

Αυτή η διαδικασία συνεχίζεται μέχρι να ολοκληρωθεί η εύρεση της πυκνά-συνδεδεμένης συστάδας. Έπειτα ένα νέο σημείο επιλέγεται και ακολουθείται η παραπάνω διαδικασία, για την εύρεση επιπλέον συστάδας ή θορύβου. Σημεία, τα οποία μάρκαρε ο αλγόριθμος ως θόρυβο και που τελικά δεν έγιναν μέρος κάποιας συστάδας αποτελούν ακραία σημεία.

Τα βασικότερα πλεονεκτήματα του αλγορίθμου DBSCAN είναι:

1. Δεν απαιτεί το εκ των προτέρων προσδιορισμό του αριθμού των συστάδων, σε αντίθεση με τον αλγόριθμο k-means.
2. Έχει τη δυνατότητα να παράγει συστάδες με αυθαίρετα σχήματα, λόγω της ύπαρξης της παραμέτρου minPts .
3. Διαθέτει καλή ευαισθησία στο θόρυβο και δεν επηρεάζεται από ακραίες τιμές.
4. Χρειάζεται μόνο δύο παραμέτρους και έχει μικρή ευαισθησία ως προς τη σειρά εμφάνισης των δεδομένων τη βάση.
5. Εφόσον τα δεδομένα έχουν γίνει κατανοητά από τον μελετητή, ο προσδιορισμός των δύο παραμέτρων δεν είναι δύσκολος.

Παρόλο που ο αλγόριθμος DBSCAN εμφανίζει αρκετά πλεονεκτήματα, διαθέτει κάποια μειονεκτήματα, εκ των οποίων τα βασικότερα είναι:

1. Τα περιθωριακά σημεία μιας συστάδας μπορούν να ανήκουν είτε σε αυτή τη συστάδα είτε σε κάποια γειτονική, ανάλογα με τη σειρά επεξεργασίας. Αυτό, ευτυχώς, δεν συμβαίνει συχνά και έχει μικρό αντίκτυπο στα αποτελέσματα.
2. Η ποιότητα των αποτελεσμάτων εξαρτάται από τη μετρική απόστασης που θα χρησιμοποιηθεί. Η πιο κοινή μετρική απόστασης είναι η Ευκλείδεια απόσταση. Ειδικά για πολυδιάστατα δεδομένα, η συγκεκριμένη μετρική είναι σχεδόν άχρηστη, λόγω της λεγόμενης «κατάρας της διαστατικότητας», κάνοντας έτσι δύσκολη την επιλογή της παραμέτρου ϵ . Ωστόσο, αυτό μπορεί να συμβεί σε οποιονδήποτε αλγόριθμο χρησιμοποιεί την Ευκλείδεια απόσταση.
3. Δεν μπορεί να συσταδοποιήσει καλά σύνολα από δεδομένα με μεγάλες διαφορές πυκνότητας, καθώς δεν μπορεί να εντοπιστεί κάποιος συνδυασμός $\text{MinPts}-\epsilon$, που να είναι κατάλληλος για όλες τις συστάδες.
4. Αν τα δεδομένα δεν έχουν γίνει κατανοητά, η επιλογή ενός κατωφλίου ϵ που να έχει νόημα μπορεί να είναι δύσκολη.

Σύμφωνα με τους (Sembiring & Zain, 2010), η διαδικασία της ομαδοποίησης (clustering analysis) περιλαμβάνει τα εξής στάδια:

- Κατανομή των δεδομένων σε ομάδες.
- Επικύρωση των αποτελεσμάτων της διαδικασίας.

Ο αλγόριθμος DBSCAN δεδομένου ότι λαμβάνει υπόψιν του την απόσταση των σημείων μεταξύ τους αλλά και την πυκνότητα των σημείων αυτών, κατά τη διάρκεια της ομαδοποίησης γίνεται ένας ταυτόχρονος έλεγχος των ομάδων που ανιχνεύονται. Οι ομάδες αυτές πρέπει να έχουν κάποια συγκεκριμένα χαρακτηριστικά, αλλιώς θεωρούνται ακρότατα (outliers).

Υπάρχουν αρκετές αναφορές για τις μετρικές με τις οποίες μπορεί να γίνει η αξιολόγηση της ομαδοποίησης. Στην παρούσα εργασία γίνεται εφαρμογή των μετρικών που παρουσιάζονται από τον (Dmitri Timašjon, 2014): κάλυψη (coverage), πυκνότητα (density) και προσαρμοστικότητα (modularity). Παρακάτω παρουσιάζονται βασικές έννοιες για τις μετρικές αυτές.

Με τον όρο κάλυψη (coverage) εννοείται η αναλογία των στοιχείων της ίδιας ομάδας σε σχέση με το συνολικό αριθμό στοιχείων στο γράφημα. Δείχνει πόσα στοιχεία βρίσκονται μέσα στην ομάδα και τον αριθμό των στοιχείων που επισημαίνονται ως θόρυβος (outliers) και ορίζεται ως:

$$coverage(C) = \frac{\text{intra-cluster edges}}{m} \quad (1)$$

Με τον όρο modularity εννοείται ένα μέτρο για τον καθορισμό της αντοχής του διαιρεμένου γραφήματος σε ενότητες. Με άλλα λόγια, ένας υψηλός δείκτης modularity σημαίνει ότι υπάρχουν περισσότερα στοιχεία μέσα σε μια ενότητα, από την αναμενόμενη πρόβλεψη να θεωρηθούν τυχαία και ορίζεται όπως παρακάτω, τα i, j ανήκουν στην ομάδα (C) που εξετάζεται.

$$modularity(C) = \sum \left(\frac{e_{ii}}{2m} - \left(\sum \frac{e_{ij}}{2m} \right)^2 \right) \quad (2)$$

Τελευταία παράμετρος αξιολόγησης είναι η πυκνότητα, όπου το c ανήκει στην ομάδα (C) που εξετάζεται.

$$\begin{aligned} density(C) = & \frac{1}{2} \left(\frac{1}{C} \sum \frac{\text{intra-cluster edges of } c}{\frac{c}{2}} \right) \\ & + \frac{1}{2} \left(1 - \frac{\text{inter-cluster edges}}{\frac{n}{2} - \sum \binom{c}{2}} \right) \end{aligned} \quad (3)$$

3.2.3 Αλγόριθμος Εξόρυξης Προτύπων Ακολουθιών

Οι αλγόριθμοι τύπου εξόρυξης πρότυπων ακολουθιών (frequent sequence mining) χρησιμοποιούνται για την εύρεση διαφορετικών μοτίβων μέσα στη βάση δεδομένων, των οποίων οι τιμές εμφανίζονται σε σειρά (sequential pattern mining) (Fournier-viger & Lin, 2017).

Τέτοιου τύπου αλγόριθμοι έχουν δημιουργηθεί κυρίως για προβλήματα όπως το σούπερ μάρκετ. Πιο συγκεκριμένα, ο αλγόριθμος ανιχνεύει “ομάδες” προϊόντων τα οποία αγοράζονται μαζί και συχνότερα από τους καταναλωτές. Για παράδειγμα, το τυρί μαζί με τα μακαρόνια. Έτσι, ο καταστηματάρχης είναι σε θέση να γνωρίζει τις προσφορές που πρόκειται να κάνει αλλά και το κέρδος που θα του αποφέρουν αυτές.

Στην παρούσα εργασία, χρησιμοποιείται για την ανίχνευση μοτίβων που παρουσιάζουν όμοια συμπεριφορά στην υπέρβαση του ορίου ταχύτητας. Σκοπός είναι να πληροφορείται ο χρήστης κατά τη διάρκεια της οδήγησης για σημεία του οδικού δικτύου στα οποία έχει παρατηρηθεί συχνή υπέρβαση του ορίου ταχύτητας από πολλούς χρήστες. Παρακάτω παρουσιάζονται βασικά στοιχεία του αλγορίθμου αυτού.

Η αρχική προσέγγιση για υλοποίηση τέτοιου είδους προβλημάτων γίνεται μέσω του αλγορίθμου Apriori, ο οποίος όμως είναι ένας αλγόριθμος εξόρυξης ακολουθιών (frequent pattern mining). Μέσω αυτού, εντοπίζονται ομάδες (patterns) αντικειμένων, τα οποία εμφανίζονται συχνά μαζί μέσα σε μια βάση δεδομένων. Πρακτικά, ο αλγόριθμος έχει ως αποτέλεσμα ομάδες αντικειμένων που ταξινομούνται σε σειρά ανάλογα με το πόσες φορές αυτά εμφανίζονται στη βάση δεδομένων. Το κυριότερο μειονέκτημα του αλγορίθμου και ταυτόχρονα ο λόγος για τον οποίο δεν χρησιμοποιήθηκε στην παρούσα εργασία είναι ότι ο εντοπισμός των αντικειμένων δεν γίνεται σε σειρά.

Ειδικότερα, στο πλαίσιο της παρούσας εργασίας, για τον εντοπισμό ομάδων με σημεία στα οποία γίνεται υπέρβαση του ορίου ταχύτητας είναι επιθυμητό η δημιουργία ομάδων να γίνει για σημεία τα οποία βρίσκονται σε σειρά πάνω σε ένα κομμάτι του οδικού δικτύου, δηλαδή το ένα πίσω από το άλλο. Ο αλγόριθμος Apriori δεν μπορεί να εξασφαλίσει το παραπάνω. Συνεπώς, χρησιμοποιείται ένας αλγόριθμος (τύπου frequent sequence mining) που τοποθετεί τα σημεία εκτός από ομάδες, όταν αυτά εμφανίζονται συχνά, αλλά και σε σειρά.

Έστω ότι ο Πίνακας 3. 1 αποτελεί αγορές καταναλωτών σε ένα κατάστημα. Η στήλη με τα προϊόντα δείχνει σειρές προϊόντων τα οποία αγοράζονται από τον εκάστοτε καταναλωτή σε διαφορετικές στιγμές. Για παράδειγμα, ένας καταναλωτής αγόρασε μαζί τα προϊόντα a και b μαζί, κάποια άλλη στιγμή το προϊόν c, τα προϊόντα f και g μαζί, το προϊόν g και τέλος το προϊόν e.

Ο αλγόριθμος ψάχνει να βρει αλληλουχίες προϊόντων, τα οποία εμφανίζονται πιο συχνά μαζί μέσα στη βάση δεδομένων που φαίνεται παρακάτω.

Πίνακας 3. 1: Παράδειγμα μορφής δεδομένων για τον αλγόριθμο εξόρυξης προτύπων ακολουθιών.

Καταναλωτής	Προϊόντα
1	{ (a,b), (c), (f,g), (g), (e) }
2	{ {a,d), (c), (b), (a,b,e,f) }
3	{ (a), (b), (f,g), (e) }
4	{ (b), (f,g) }

Προκειμένου να γίνει αυτό, τίθεται ένα όριο από τον χρήστη (support). Η παράμετρος αυτή υποδεικνύει, τον ελάχιστο αριθμό αλληλουχιών, όπου ένα μοτίβο πρέπει να εμφανίζεται για να θεωρηθεί συχνά εμφανιζόμενο. Για παράδειγμα, εάν αυτό το όριο τεθεί ίσο με δύο, εντοπίζονται όλες οι υπό - αλληλουχίες προϊόντων, οι οποίες εμφανίζονται σε τουλάχιστον δύο αλληλουχίες στην αρχική βάση δεδομένων. Στο συγκεκριμένο παράδειγμα, βρέθηκαν 29 υπό - αλληλουχίες που συμπίπτουν με αυτό το όριο και φαίνονται στο Πίνακα 3. 2. Η πρώτη στήλη δείχνει τα μοτίβα που εντοπίζονται και η δεύτερη δείχνει το παραπάνω όριο (support) για το κάθε μοτίβο.

Πίνακας 3. 2: Αλληλουχίες προϊόντων και συχνότητα εμφάνισης.

Pattern	Sup.
$\langle \{a\} \rangle$	3
$\langle \{a\}, \{g\} \rangle$	2
$\langle \{a\}, \{g\}, \{e\} \rangle$	2
$\langle \{a\}, \{f\} \rangle$	3
$\langle \{a\}, \{f\}, \{e\} \rangle$	2
$\langle \{a\}, \{c\} \rangle$	2
$\langle \{a\}, \{c\}, \{f\} \rangle$	2
$\langle \{a\}, \{c\}, \{e\} \rangle$	2
$\langle \{a\}, \{b\} \rangle$	2
$\langle \{a\}, \{b\}, \{f\} \rangle$	2
$\langle \{a\}, \{b\}, \{e\} \rangle$	2
$\langle \{a\}, \{e\} \rangle$	3
$\langle \{a, b\} \rangle$	2
$\langle \{b\} \rangle$	4
$\langle \{b\}, \{g\} \rangle$	3
$\langle \{b\}, \{g\}, \{e\} \rangle$	2
$\langle \{b\}, \{f\} \rangle$	4
$\langle \{b\}, \{f, g\} \rangle$	3
$\langle \{b\}, \{f\}, \{e\} \rangle$	2
$\langle \{b\}, \{e\} \rangle$	3
$\langle \{c\} \rangle$	2
$\langle \{c\}, \{f\} \rangle$	2
$\langle \{c\}, \{e\} \rangle$	2
$\langle \{e\} \rangle$	3
$\langle \{f\} \rangle$	4
$\langle \{f, g\} \rangle$	3
$\langle \{f\}, \{e\} \rangle$	2
$\langle \{g\} \rangle$	3
$\langle \{g\}, \{e\} \rangle$	2

Για να γίνει πιο κατανοητό, το προϊόν a εμφανίζεται σε τρεις αλληλουχίες της αρχικής βάσης δεδομένων, τα προϊόντα a και g εμφανίζονται δύο φορές μαζί.

Για την αξιολόγηση, λοιπόν, του προβλήματος αυτού, χρησιμοποιείται το πακέτο Rymining της γλώσσας προγραμματισμού Python. Το πακέτο αυτό περιλαμβάνει αρκετούς τύπους αλγορίθμων για την ανάλυση δεδομένων. Στην παρούσα εργασία, χρησιμοποιείται ο αλγόριθμος seqmining.

4. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1 ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Για τη συλλογή των δεδομένων σε κίνηση με μηχανοκίνητο όχημα χρησιμοποιήθηκε η εφαρμογή OSeven Telematics (www.oseven.io). Πρόκειται για μια πρόσφατη καινοτομία που χρησιμοποιείται σε επίπεδο έρευνας και κατέστη δυνατή χάρη στις τελευταίες εξελίξεις της κινητής τηλεφωνίας. Η εφαρμογή αυτή έχει το πλεονέκτημα ότι μπορεί να λειτουργεί χωρίς να επιβαρύνει τα χρήστη, καθώς ενεργοποιείται και απενεργοποιείται χωρίς την παρέμβαση του οδηγού και έχει τη δυνατότητα συλλογής μεγάλου όγκου δεδομένων, μέσω των οποίων προκύπτουν πολύτιμα μεταδεδομένα (metadata) για τη συμπεριφορά του οδηγού. Μοναδική προϋπόθεση είναι ο οδηγός να έχει εγκαταστήσει την εφαρμογή στο κινητό του και να έχει εισέλθει σε αυτή με το όνομα χρήστη και το κωδικό του. Η εφαρμογή δίνει στον οδηγό τη δυνατότητα να πραγματοποιήσει στάσεις λίγων λεπτών, χωρίς να θεωρηθεί ότι ο οδηγός έχει ολοκληρώσει την εκάστοτε διαδρομή.

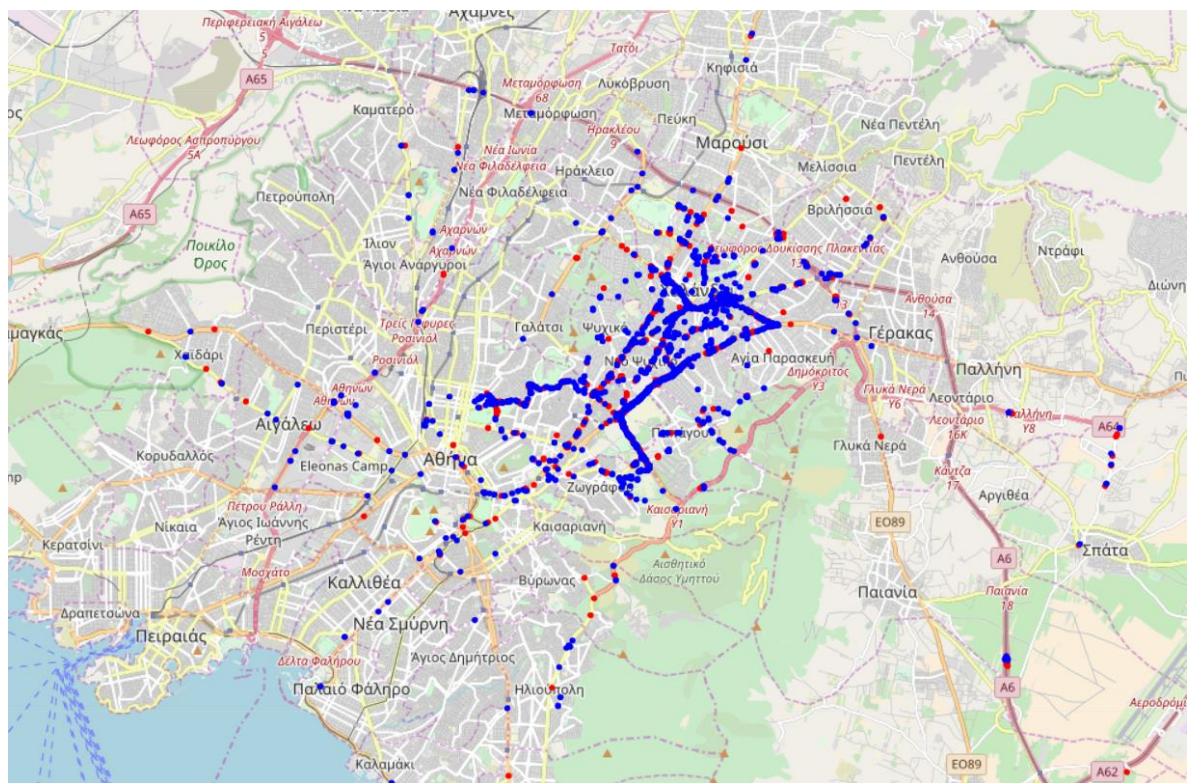
Για κάθε ταξίδι, η επίδοση του οδηγού αναλύεται μέσα από μια σειρά παραμέτρων που προκύπτουν ύστερα από ανάλυση χρονοσειρών από τους αισθητήρες του κινητού (επιταχυνσιόμετρο, γυροσκόπιο κτλ).

Τα διαθέσιμα δεδομένα περιείχαν τις εξής πληροφορίες (μεταβλητές) σε επίπεδο ταξιδιού:

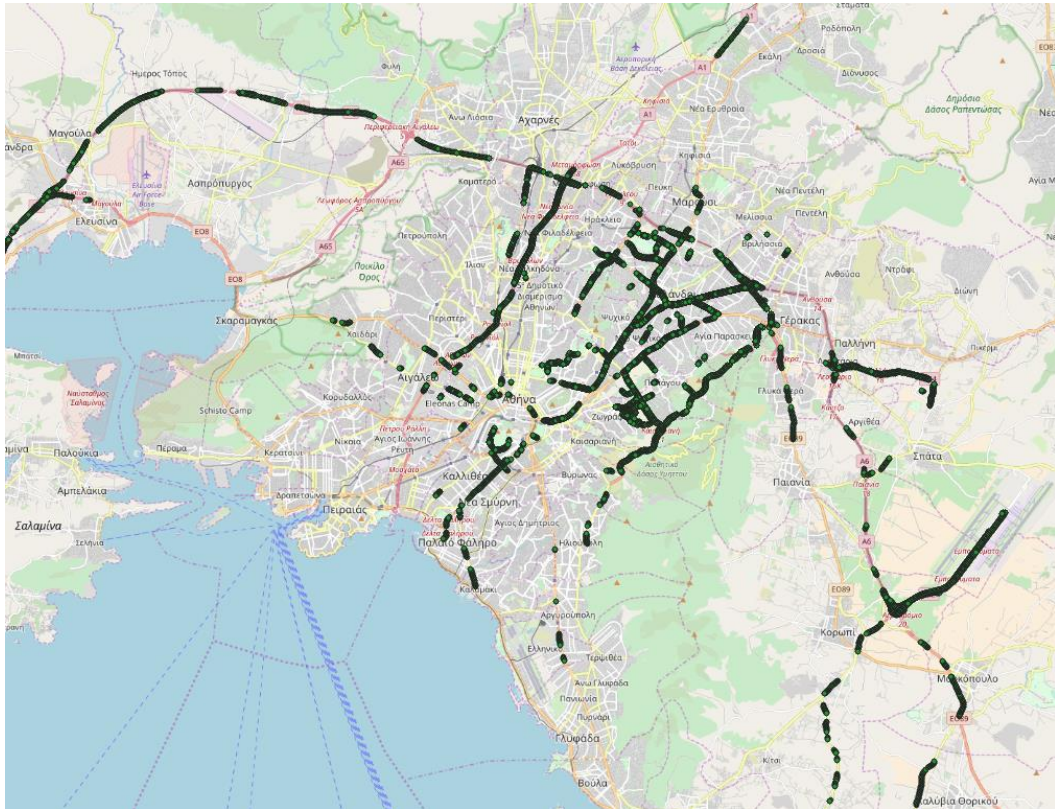
- Αριθμός απότομων επιταχύνσεων (HA)
- Αριθμός απότομων επιβραδύνσεων (HB)
- Υπέρβαση ορίου ταχύτητας (Speeding)

Οι μεταβλητές αυτές περιγράφουν την επιθετικότητα της οδήγησης (aggressiveness), την έλλειψη προσοχής (anticipation), την απόσπαση προσοχής (distraction), και την επικίνδυνη οδήγηση (risky driving). Με βάση αυτές τις μεταβλητές θα γίνει η ανάλυση και η ανίχνευση των συστάδων επικινδυνότητας.

Χρησιμοποιήθηκαν δύο βάσεις δεδομένων. Η πρώτη επεξεργασία πραγματοποιείται για τις κινήσεις ενός μόνο οδηγού για 10142 ταξίδια. Για αυτή τη βάση δεδομένων, πραγματοποιούνται συνολικά από το χρήστη 3079 απότομες επιταχύνσεις (ha events) και 1475 απότομα φρεναρίσματα (hb events), τα οποία απεικονίζονται στον χάρτη της Εικόνας 4. 1. Με μπλε χρώμα απεικονίζονται οι απότομες επιταχύνσεις που πραγματοποιούνται από τον οδηγό και με κόκκινο οι αντίστοιχες απότομες επιβραδύνσεις. Στην Εικόνα 4. 2 απεικονίζονται όλα τα σημεία στα οποία παρατηρήθηκε υπέρβαση του ορίου ταχύτητας από το χρήστη.

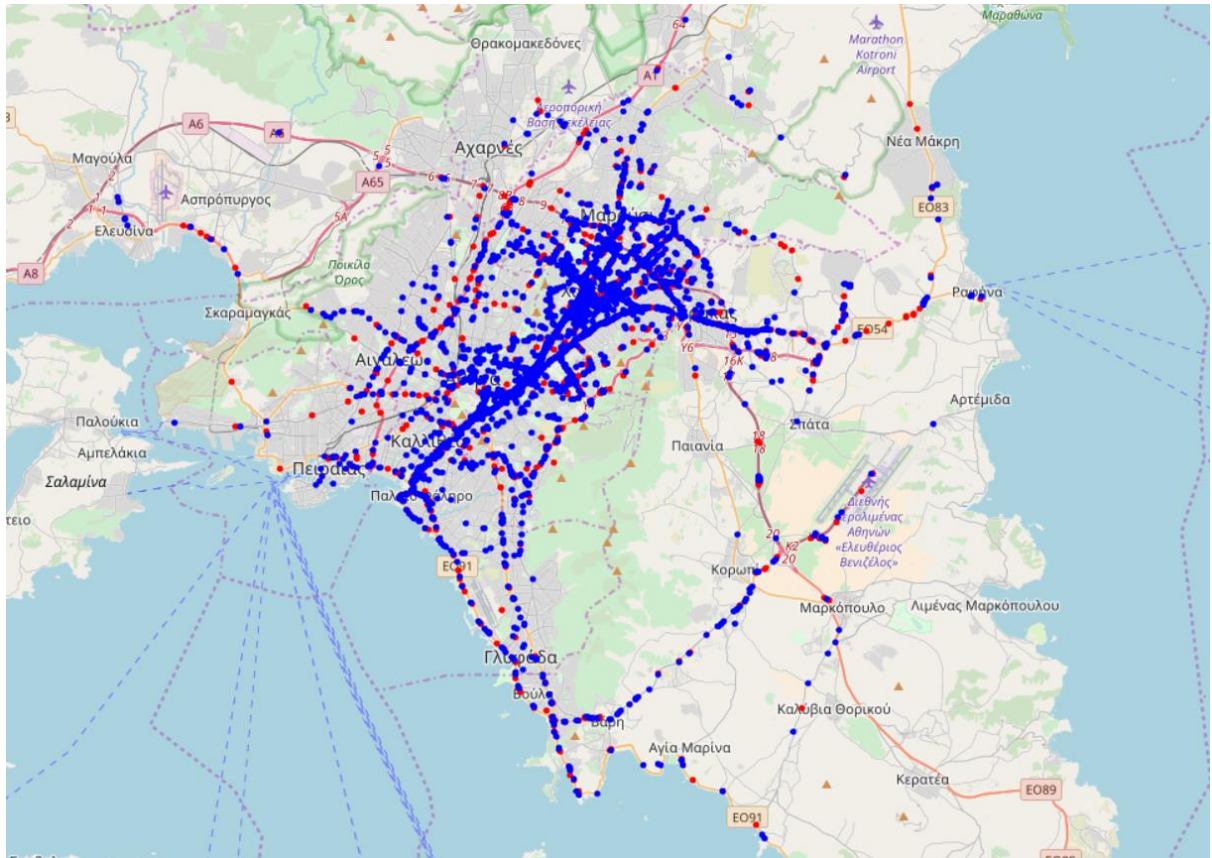


Εικόνα 4. 1: Απεικόνιση του συνόλου των σημειακών συμβάντων μεμονωμένου χρήστη.

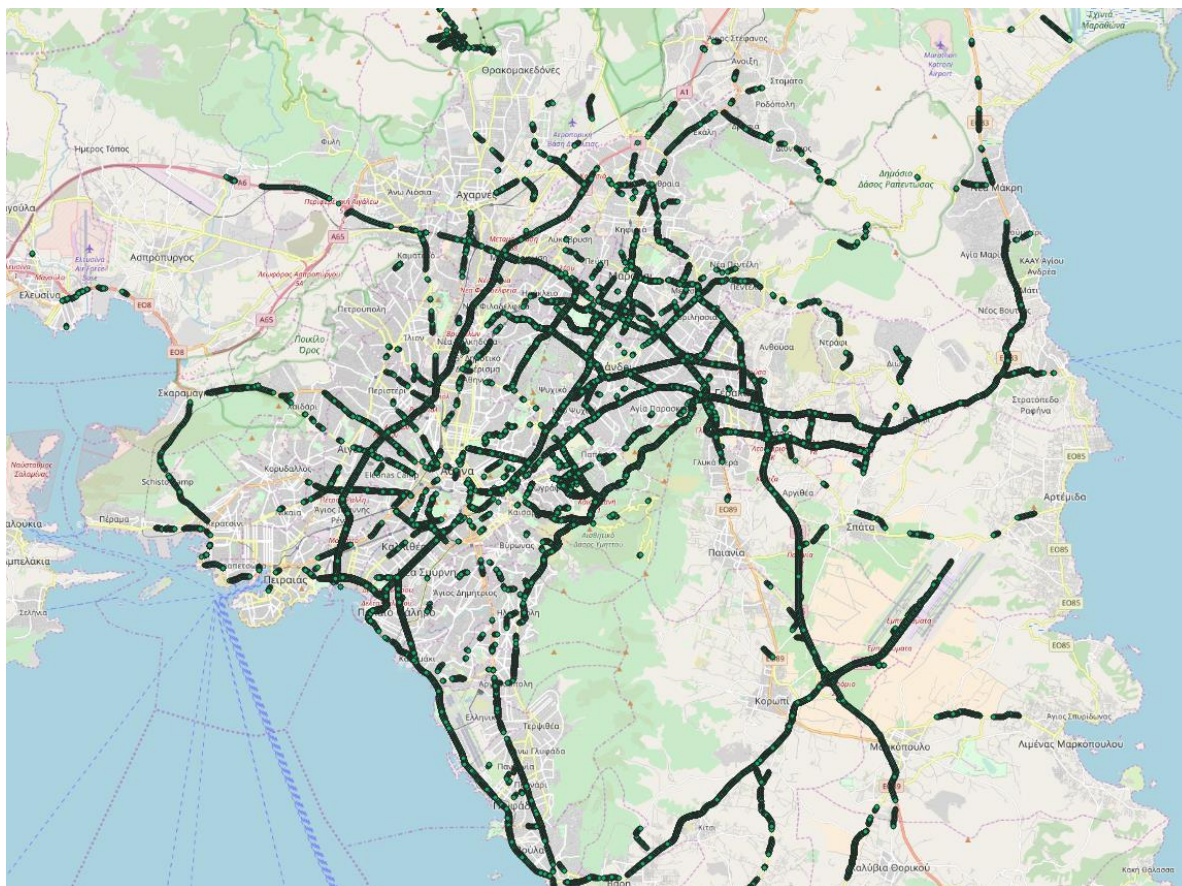


Εικόνα 4. 2: Απεικόνιση του συνόλου των συμβάντων υπέρβασης ορίου ταχύτητας μεμονωμένου χρήστη.

Η διαθέσιμη βάση δεδομένων της επόμενης φάσης, αποτελείται από 3262 ταξίδια, μέσα στο οδικό δίκτυο της Αθήνας, από 147 τυχαίους οδηγούς ηλικίας κατά κύριο λόγο από 25 έως 55 ετών με τους άνδρες και τις γυναίκες να καταλαμβάνουν περίπου το 65% και 35% του δείγματος αντίστοιχα. Για τη βάση δεδομένων από διάφορους οδηγούς, έχουμε συνολικά 4894 απότομες επιταχύνσεις (ha events) και 1981 απότομα φρεναρίσματα (hb events), όπως φαίνεται στον χάρτη της Εικόνας 4. 3. Με μπλε χρώμα απεικονίζονται οι απότομες επιταχύνσεις που πραγματοποιούνται από τους οδηγούς και με κόκκινο οι αντίστοιχες απότομες επιβραδύνσεις. Στον χάρτη της Εικόνας 4. 4 απεικονίζονται όλα τα σημεία στα οποία παρατηρήθηκε υπέρβαση του ορίου ταχύτητας.



Εικόνα 4. 3: Απεικόνιση του συνόλου των σημειακών συμβάντων πολλών χρηστών.



Εικόνα 4. 4: Απεικόνιση του συνόλου των συμβάντων υπέρβασης ορίου ταχύτητας πολλών χρηστών.

Συγκεντρωτικά, ο Πίνακας 4. 1 συνοψίζει τα παραπάνω για τις δυο βάσεις δεδομένων που χρησιμοποιούνται και έτσι έχουμε:

Πίνακας 4. 1: Δεδομένα προς επεξεργασία.

Πλήθος χρηστών	Πλήθος ταξιδιών	Σύνολο απότομων επιταχύνσεων	Σύνολο απότομων επιβραδύνσεων
1	10142	3079	1475
147	3262	4894	1981

4.2 ΑΝΑΛΥΣΗ ΑΠΟΤΟΜΩΝ ΕΠΙΤΑΧΥΝΣΕΩΝ ΚΑΙ ΑΠΟΤΟΜΩΝ ΕΠΙΒΡΑΔΥΝΣΕΩΝ

Τα σημεία όπου πραγματοποιούνται απότομες επιταχύνσεις ή επιβραδύνσεις μπορούν να αναλυθούν με τον ίδιο τρόπο, δηλαδή με την υλοποίηση του ίδιου αλγορίθμου. Για το λόγο αυτό, στήνεται ένας αλγόριθμος DBSCAN για την επεξεργασία τους.

4.2.1 Προκαταρκτική Επεξεργασία Δεδομένων

Επόμενο στάδιο από τη συλλογή δεδομένων είναι η επεξεργασία αυτών με τρόπο τέτοιο ώστε να μπορούν να εισαχθούν στους αλγορίθμους.

Στην περίπτωση αυτή, η μορφοποίηση των δεδομένων παίζει σημαντικό ρόλο. Έτσι, σε πρώτη φάση, από την αρχική βάση δεδομένων απαραίτητα προς ανάλυση αποτελούν τα σημεία όπου ο παρατηρήθηκε κάποιο συμβάν για το χρήστη. Επομένως, ενδιαφέρουν σε πρώτο στάδιο οι συντεταγμένες των σημείων αυτών (γεωγραφικό μήκος και γεωγραφικό πλάτος). Τα δεδομένα τα οποία εισέρχονται στον αλγόριθμο έχουν την παρακάτω μορφή:

Πίνακας 4. 2: Δεδομένα προς εισαγωγή στον αλγόριθμο DBSCAN.

location Latitude	location Longitude
38.04670861	23.83951608
38.04005392	23.83042967
38.03375949	23.82241076
38.00688791	23.79973160
38.02590809	23.78588585
37.96658278	23.77001755
37.98392745	23.78310937
38.01831900	23.78776643

4.2.2 Ανάλυση Συμπεριφοράς Μεμονωμένου Χρήστη

Για την ανάλυση της συμπεριφοράς μεμονωμένου χρήστη, απομονώνονται και αναλύονται τα ταξίδια ενός τυχαίου χρήστη. Ο χρήστης αυτός επιλέχθηκε με κριτήριο τον αριθμό ταξιδιών εντός του οδικού δικτύου της Αθήνας, ώστε να προκύψουν αξιόπιστα αποτελέσματα.

Η ανάλυση βασίζεται στη συσταδοποίηση της οδηγικής συμπεριφοράς που περιγράφεται από τις μεταβλητές HA και HB με εφαρμογή του αλγορίθμου DBSCAN.

Ο αλγόριθμος απαιτεί τον καθορισμό δύο παραμέτρων προκειμένου να γίνει η ένταξη των σημείων σε συστάδες, την ακτίνα ϵ και του αριθμού των σημείων που ορίζουν μια συστάδα N .

Για την ακτίνα της κάθε συστάδας επιλέγεται η τιμή των 100 μέτρων, όπου σε επίπεδο οδικού δικτύου της Αθήνας κρίθηκε ικανοποιητική. Η τιμή αυτή κρίθηκε ικανοποιητική δεδομένου ότι η ανάλυση των παρεχόμενων βάσεων δεδομένων βρίσκεται στο οδικό δίκτυο της Αθήνας και όχι σε κάποιο μεγάλο αστικό κέντρο ή στο οδικό δίκτυο μιας ολόκληρης χώρας. Η παράμετρος αυτή είναι κοινή σε επίπεδο ενός χρήστη με την αντίστοιχη σε επίπεδο διαφορετικών χρηστών σε συγκεκριμένο αριθμό ταξιδιών.

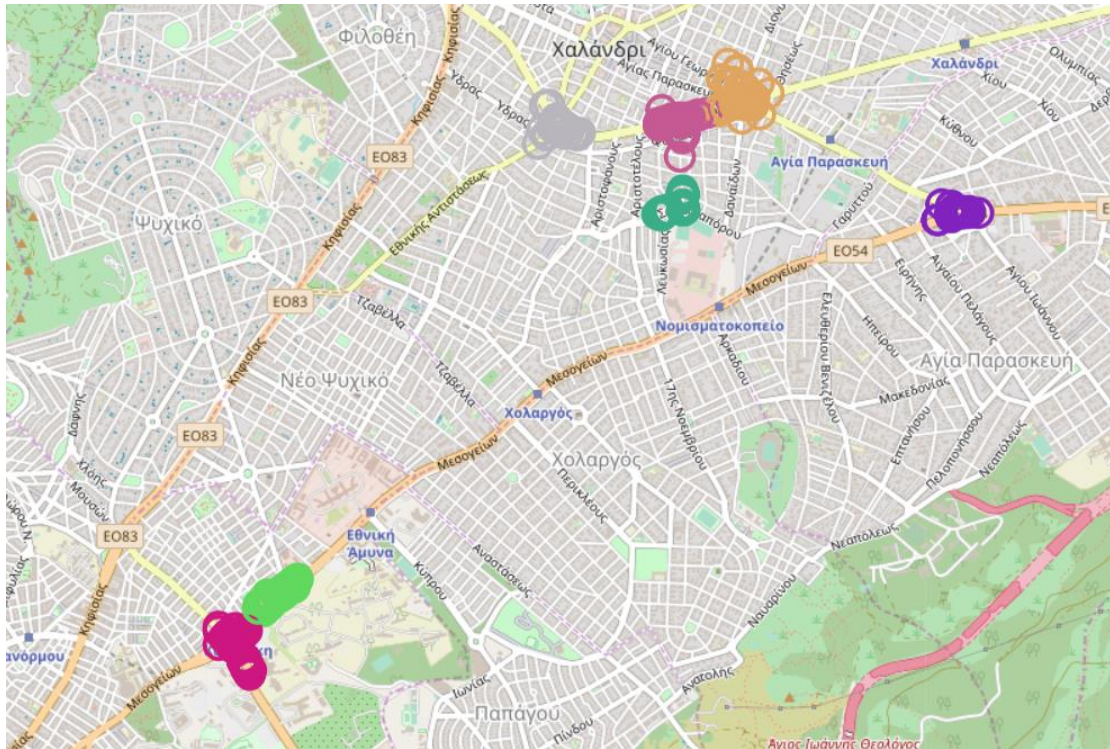
Για τον προσδιορισμό του αριθμού των σημείων που ορίζουν μια συστάδα σε επίπεδο ενός χρήστη ακολουθείται μια επαναληπτική διαδικασία εκτίμησης της ευαισθησίας του αποτελέσματος της ομαδοποίησης (αριθμός ομάδων) σε σχέση με διαφορετικές τιμές της παραμέτρου N και με τη θεώρηση ότι η ακτίνα επιρροής μιας ομάδας παραμένει σταθερή και ίση με 100 μέτρα.

Από την ανάλυση ευαισθησίας προκύπτει ότι για $N = 40$ σημεία για συμβάντα απότομων επιταχύνσεων, ο αλγόριθμος εντοπίζει οκτώ διαφορετικές ομάδες.

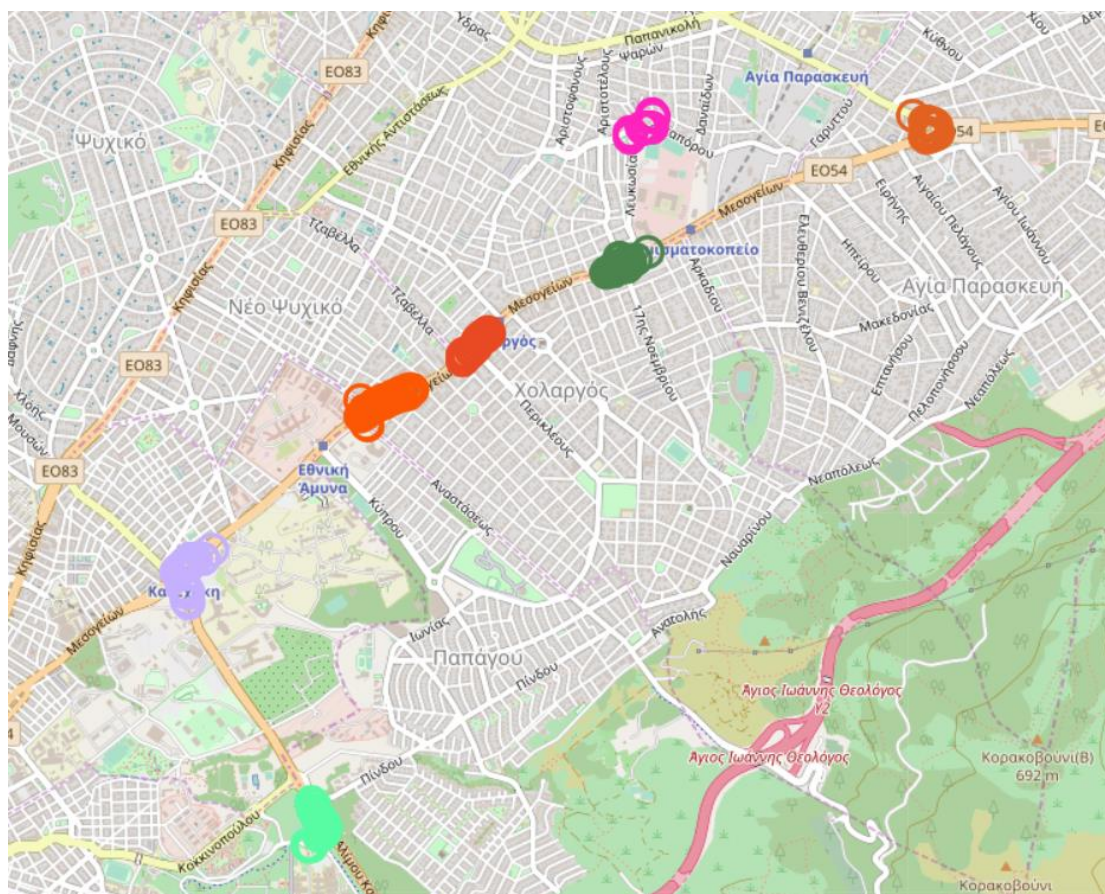
Για συμβάντα απότομων επιβραδύνσεων, κάθε συστάδα αποτελείται από 30 το πολύ σημεία και προκύπτουν επτά διαφορετικές ομάδες.

Τα αποτελέσματα οπτικοποιούνται σε χάρτες επικινδυνότητας που φαίνονται στις Εικόνες 4. 5 και 4. 6.

Για κάθε συστάδα, επιλέγεται ένα διαφορετικό χρώμα προκειμένου να γίνει η απεικόνιση της στο χάρτη, όπως φαίνεται και στη παρακάτω εικόνα. Κάθε συστάδα αποτελείται από διαφορετικά σημεία στα οποία παρατηρήθηκε ένα συμβάν, καθένα από τα οποία απεικονίζεται με έναν κύκλο.



Εικόνα 4. 5: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιταχύνσεων μεμονωμένου χρήστη.



Εικόνα 4. 6: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιβραδύνσεων μεμονωμένου χρήστη.

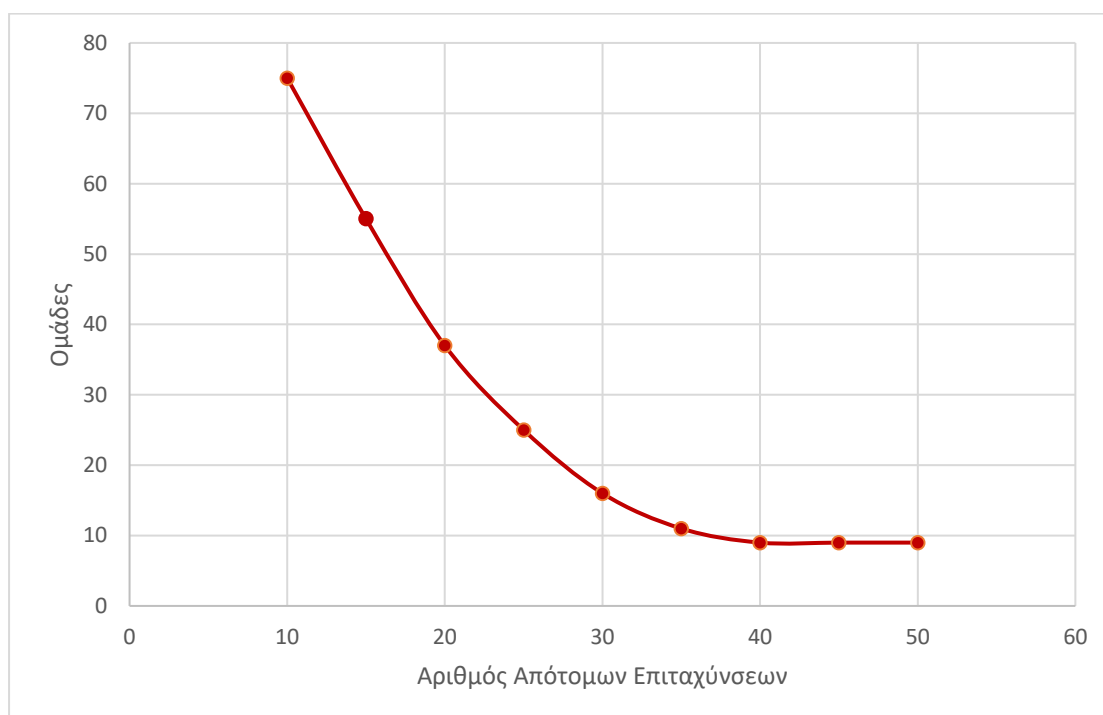
4.2.3 Ανάλυση Συμπεριφοράς Πολλών Χρηστών

Αντίστοιχη εφαρμογή με αυτή σε επίπεδο ενός χρήστη πραγματοποιείται και στην περίπτωση περισσότερων 147 τυχαίων οδηγών εντός του οδικού δικτύου της Αθήνας για συγκεκριμένο αριθμό ταξιδιών.

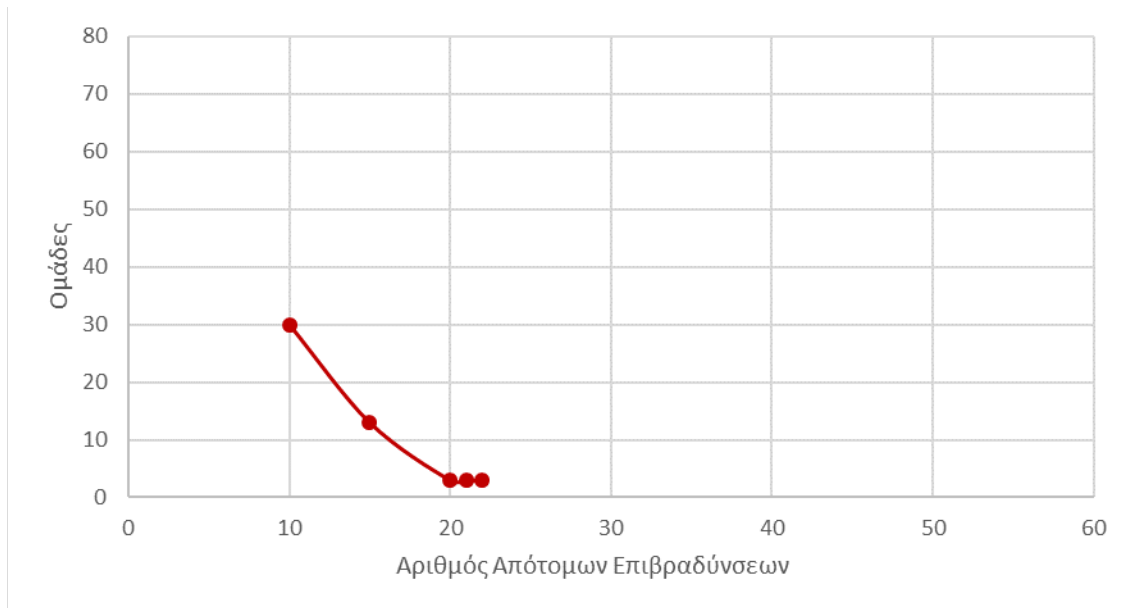
Στον πίνακα 4. 3, φαίνονται οι διαφορετικές ομάδες που έχουν δημιουργηθεί από την εφαρμογή του αλγορίθμου στη βάση δεδομένων για απότομες επιταχύνσεις και απότομες επιβραδύνσεις, αναλόγως τον μέγιστο αριθμό των σημείων που θα περιλαμβάνει κάθε συστάδα. Τα αποτελέσματα οπτικοποιούνται στα Διαγράμματα 4. 1 και 4. 2.

Πίνακας 4. 3: Ομάδες επικινδυνότητας σε σχέση με τον αριθμό των συμβάντων.

Απότομες Επιταχύνσεις		Απότομες Επιβραδύνσεις	
N	Ομάδες	N	Ομάδες
10	75	10	30
15	55	15	13
20	37	20	3
25	25	21	3
30	11	22	3
35	16		
40	9		
45	9		
50	9		



Διάγραμμα 4. 1: Σχέση ομάδων επικινδυνότητας και αριθμού σημείων απότομων επιταχύνσεων κάθε ομάδας.



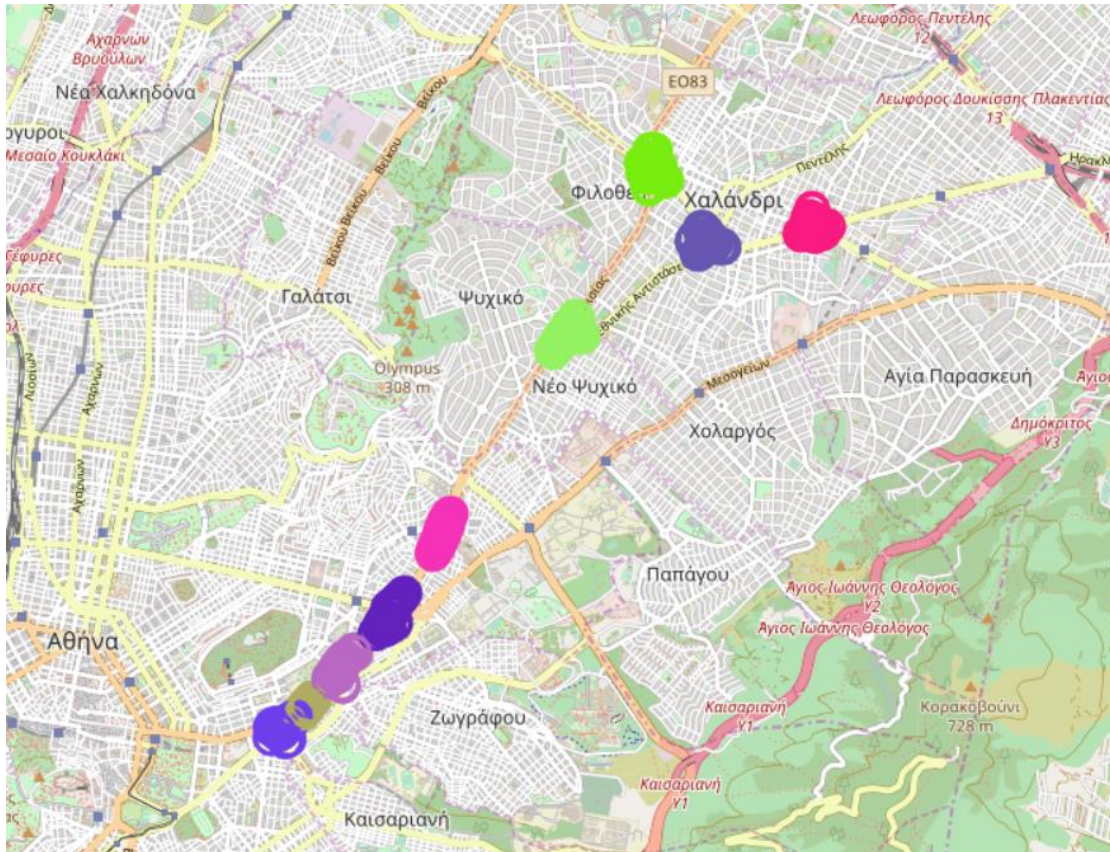
Διάγραμμα 4. 2: Σχέση ομάδων επικινδυνότητας και αριθμού σημείων απότομων επιβραδύνσεων κάθε ομάδας.

Επομένως, για τον εντοπισμό των συστάδων για συμβάντα απότομων επιταχύνσεων, επιλέγεται ο αριθμός 40, ως η τιμή της παραμέτρου για τον αριθμό των σημείων που ορίζουν μια συστάδα, ενώ ο αριθμός των ομάδων σταθεροποιείται για την τιμή της παραμέτρου ίση με 20 σημεία, για τις απότομες επιβραδύνσεις.

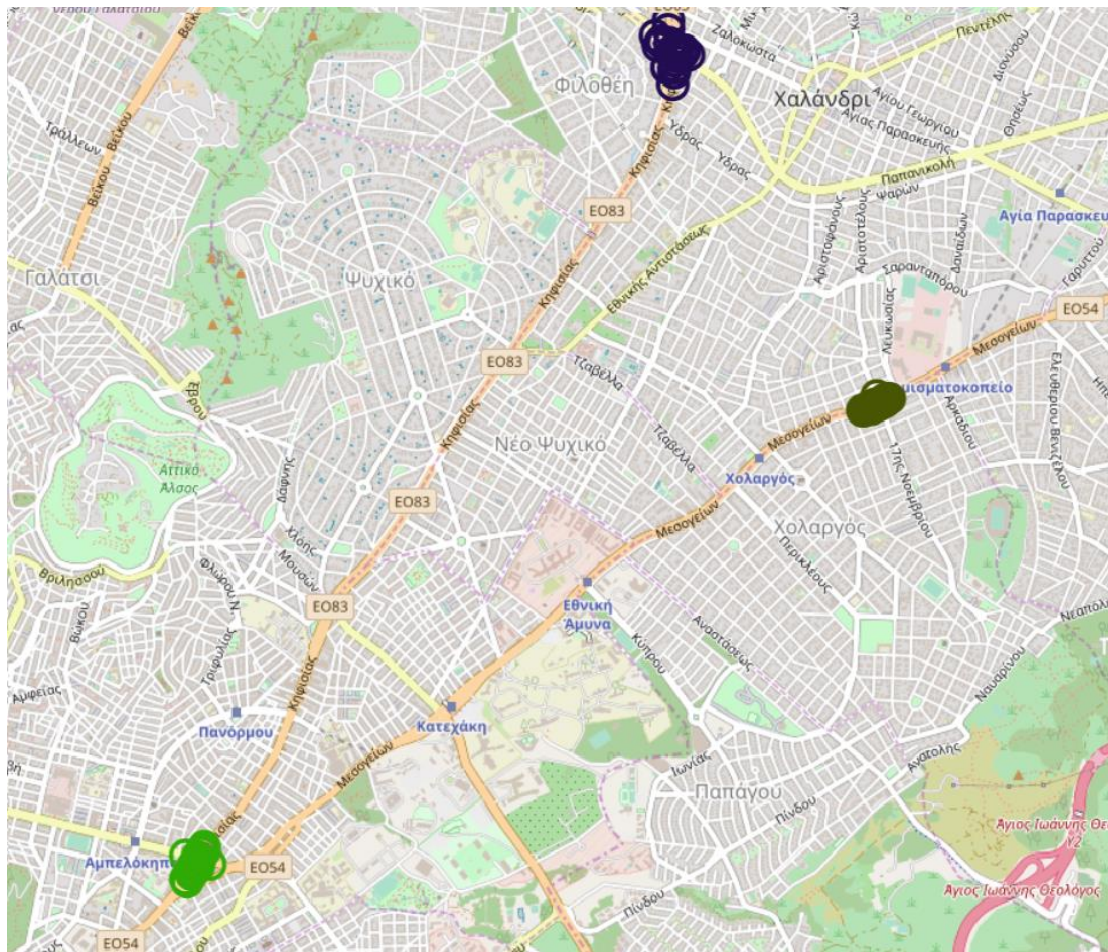
Αξίζει εδώ να σημειωθεί πως όσο ο αριθμός αυτών των σημείων γίνεται και μικρότερος, μετά από κάποιο σημείο ο αλγόριθμος θα εντοπίσει μια και μοναδική συστάδα, μέχρις ότου τελικά να μην ανιχνεύει καμία (Number of Clusters = 0). Κάτι τέτοιο ήταν αναμενόμενο να συμβεί, καθώς όταν τεθεί μεγάλο όριο για τον αριθμό των σημείων που ορίζουν μια ομάδα, δεν παρατηρείται ομοιομορφία για τα σημεία της βάσης δεδομένων.

Οι κρίσιμες θέσεις με βάση την παραπάνω ανάλυση με βάση τις απότομες επιταχύνσεις και επιβραδύνσεις φαίνονται στις Εικόνες 4. 7 και 4. 8 αντίστοιχα. Γενικότερα, κατά την ανάλυση των παρεχόμενων δεδομένων παρατηρούνται τα παρακάτω για τη βάση δεδομένων από διαφορετικούς χρήστες:

- Ο συνολικός αριθμός των απότομων επιταχύνσεων στη βάση δεδομένων είναι ίσος με 4894. Τα 4176 από αυτά, ο αλγόριθμος θεωρεί ως “θόρυβο” (outliers), ένα ποσοστό ίσο με 85.39% επί των συνολικών.
- Αντίστοιχα, ο συνολικός αριθμός των απότομων επιβραδύνσεων στη βάση δεδομένων είναι ίσος με 1981, εκ των οποίων τα 1907 επισημαίνονται ως outliers, ποσοστό ίσο με 92.27%.



Εικόνα 4. 7: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιταχύνσεων πολλών χρηστών.



Εικόνα 4. 8: Χάρτης Επικινδυνότητας για συμβάντα απότομων επιβραδύνσεων πολλών χρηστών.

Συνεπώς, αρκετά μεγάλο ποσοστό των σημείων όπου οι οδηγοί πραγματοποιούν είτε απότομες επιταχύνσεις είτε απότομες επιβραδύνσεις, ο αλγόριθμος τα θεωρεί ως outliers. Αυτό μπορεί να εξηγηθεί με διάφορους τρόπους. Αρχικά, τα σημεία αυτά δεν εμφανίζουν καμία ομοιομορφία στη συμπεριφορά τους στο οδικό δίκτυο. Επιπλέον, τα σημεία αυτά είναι πιθανό να απέχουν μεταξύ τους μεγαλύτερη απόσταση από αυτήν που έχει τεθεί για τη δημιουργία μιας συστάδας ή αν έχουν την επιτρεπτή απόσταση μεταξύ τους, δεν ικανοποιούν το όριο για το πλήθος των σημείων που ορίζουν μια συστάδα.

Σημαντικό επίσης είναι να αναλυθεί η ταχύτητα στην οποία πραγματοποιούνται οι απότομες επιταχύνσεις και επιβραδύνσεις που εντάχθηκαν σε κάποια από τις ομάδες του αλγορίθμου. Συνεπώς προκύπτει:

- Για τις απότομες επιταχύνσεις, ένα ποσοστό ίσο με 12,53% (545 από τις συνολικές) επί των συνολικών επιταχύνσεων που εντάσσονται σε μια συστάδα, πραγματοποιείται σε ταχύτητα μεγαλύτερη από 30 km/h.
- Για τις απότομες επιβραδύνσεις, ένα ποσοστό ίσο με 17,36% (293 από τις συνολικές) επί των συνολικών επιβραδύνσεων που εντάσσονται σε μια συστάδα, πραγματοποιείται σε ταχύτητα μεγαλύτερη από 30 km/h.

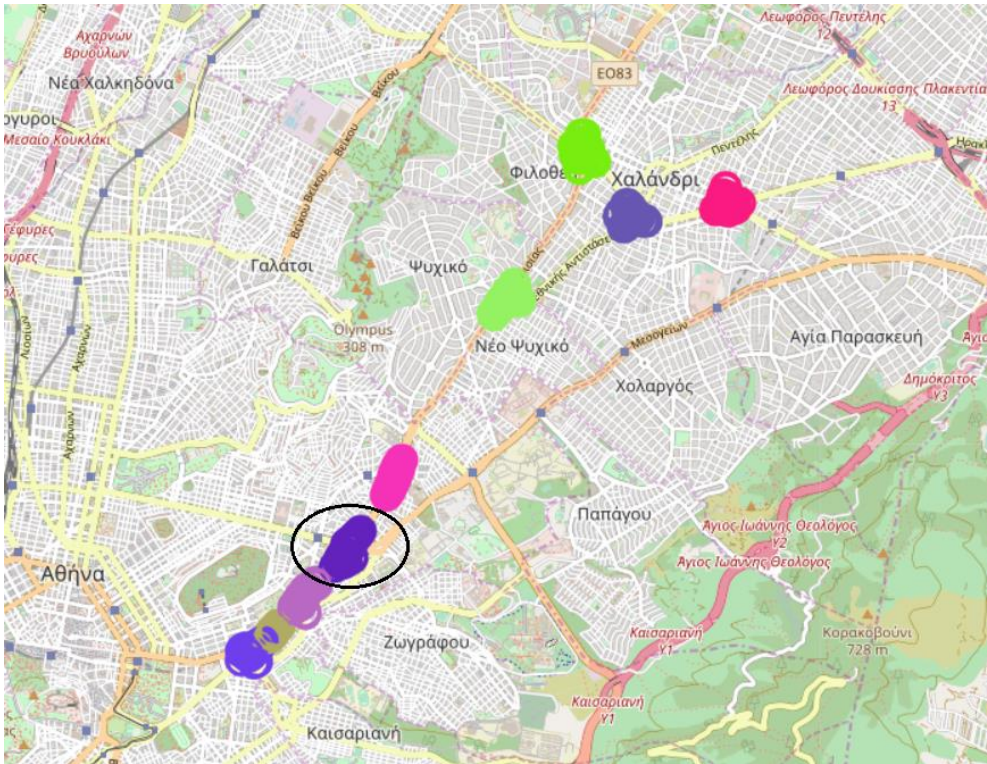
Από τις ταχύτητες που πραγματοποιούνται τα παραπάνω, είναι φανερό πως οι περισσότερες απότομες επιταχύνσεις και φρεναρίσματα πραγματοποιούνται σε μικρότερες ταχύτητες, πράγμα το οποίο επιβεβαιώνει το γεγονός πως τα περισσότερα από αυτά τα events γίνονται κοντά σε φωτεινό σηματοδότη ή σε ειδική σήμανση.

Στη συνέχεια, κρίθηκε σκόπιμο να βρεθούν οι πιο επικίνδυνες μέσα στο οδικό δίκτυο. Για το σκοπό αυτό, για κάθε μια από τις κατηγορίες των γεγονότων που πραγματοποιούνται βρίσκονται οι πιο κρίσιμες σύμφωνα με τον αριθμό των σημείων που περιλαμβάνει η κάθε συστάδα. Έτσι, για τις απότομες επιταχύνσεις οι πιο επικίνδυνες ομάδες είναι οι εξής:

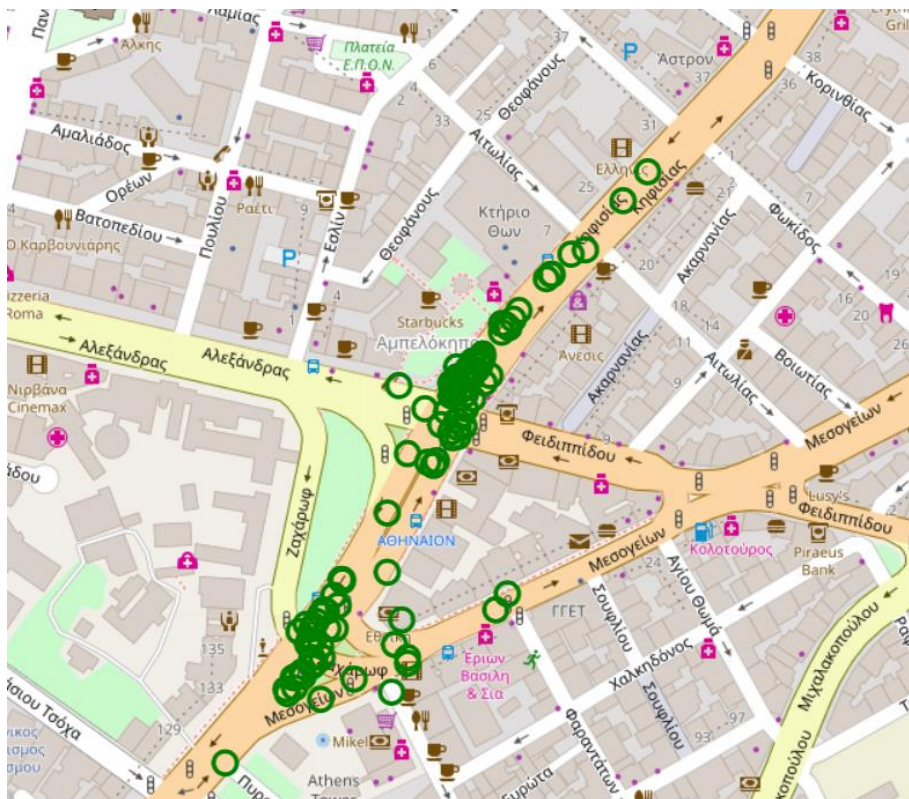
Πίνακας 4. 4: Επικίνδυνες ομάδες για συμβάντα απότομων επιταχύνσεων.

Ομάδα	Αριθμός Σημείων
0	105
3	91
5	91

Η ομάδα, λοιπόν, με τον αριθμό 0 εμφανίζεται ως η πιο κρίσιμη. Βρίσκεται στη διασταύρωση των Λεωφόρων Κηφισίας και Αλεξάνδρας και φαίνεται στον επόμενο χάρτη. Πιο συγκεκριμένα, από τη συστάδα αυτή παρατηρούνται συνολικά 105 διελεύσεις από 27 διαφορετικούς οδηγούς. Οι 13 από τις διελεύσεις αυτές πραγματοποιούνται με ταχύτητα μεγαλύτερη από 30 km/h. Σε κανένα από τα σημεία της συστάδας αυτής δεν παρατηρείται υπέρβαση του ορίου ταχύτητας.

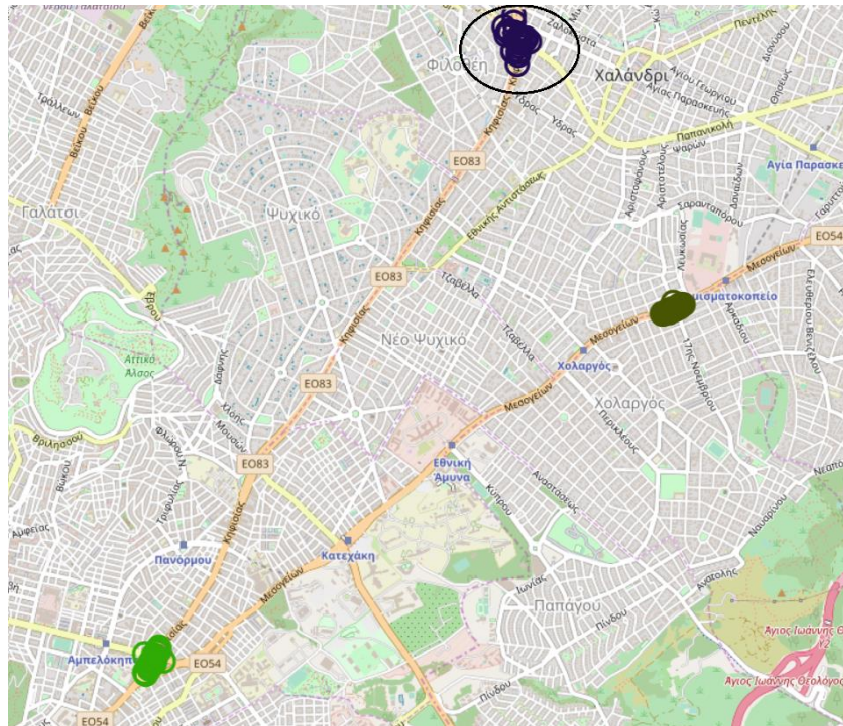


Εικόνα 4. 9: Κρίσιμη ομάδα για συμβάντα απότομων επιταχύνσεων.



Εικόνα 4. 10: Κρίσιμος κόμβος για συμβάντα απότομων επιταχύνσεων.

Αντίστοιχα, για τις απότομες επιβραδύνσεις, ως πιο επικίνδυνη χαρακτηρίστηκε η συστάδα με τον αριθμό 2, η οποία περιλαμβάνει 29 συνολικά σημεία. Βρίσκεται στη διασταύρωση της Λεωφόρου Κηφισίας με την οδό Καποδιστρίου. Από την ομάδα αυτή, παρατηρούνται συνολικά 29 διελεύσεις από 16 διαφορετικούς οδηγούς. Οι 25 από τις διελεύσεις αυτές πραγματοποιούνται με ταχύτητα μεγαλύτερη από 30 Km/h. Μόνο σε ένα σημείο παρατηρήθηκε υπέρβαση του ορίου ταχύτητας ίση με 11,66%, για περιοχή με όριο ταχύτητας τα 50 km/h.



Εικόνα 4. 11: Κρίσιμη ομάδα για συμβάντα απότομων επιβραδύνσεων.

Παρατηρείται πως στα σημεία όπου γίνεται μια απότομη επιτάχυνση δεν γίνεται, κατά κανόνα, απότομη επιβράδυνση στην επόμενη φάση. Επομένως, παρατηρείται μη αναμενόμενη συμπεριφορά των οδηγών. Αυτό διακρίνεται και από το γεγονός ότι έχουμε διαφοροποίηση στους οδηγούς για διέλευση από ένα συγκεκριμένο σημείο. Συνεπώς, είναι φανερό πως υπάρχει ομοιομορφία στη συμπεριφορά των χρηστών στο οδικό δίκτυο, δεδομένου ότι οδηγούν με τον ίδιο τρόπο διαφορετικοί οδηγοί σε συγκεκριμένο τμήμα του οδικού δικτύου.

Όσον αφορά τη ταχύτητα διέλευσης από αυτά τα σημεία, για τις απότομες επιταχύνσεις, παρατηρείται πως το μεγαλύτερο ποσοστό αυτών γίνεται σε μικρές ταχύτητες. Συνδέεται λοιπόν, με την εκκίνηση του οχήματος.

Για τις απότομες επιβραδύνσεις, παρατηρήθηκε πως πραγματοποιούνται, κατά κανόνα, σε μεγαλύτερες ταχύτητες. Αν τεθεί ένα αυξημένο όριο, για παράδειγμα τα 50 km/h, ο αριθμός των απότομων επιβραδύνσεων από 25 που ορίστηκε για το όριο των 30 km/h έγινε 24. Δεν μειώθηκε σημαντικά. Άρα, παρατηρείται πως οι περισσότερες απότομες επιβραδύνσεις γίνονται έπειτα από μεγάλη ταχύτητα οδήγησης.

4.3 ΑΝΑΛΥΣΗ ΣΥΜΒΑΝΤΩΝ ΥΠΕΡΒΑΣΗΣ ΟΡΙΟΥ ΤΑΧΥΤΗΤΑΣ

Όπως αναφέρθηκε και στο Κεφάλαιο 3, για την ανάλυση των γεγονότων, όπως στην περίπτωση αυτή, σημείων στα οποία πραγματοποιείται υπέρβαση του ορίου ταχύτητας, χρησιμοποιείται ένας αλγόριθμος τύπου frequent sequence mining.

4.3.1 Προκαταρκτική Επεξεργασία Δεδομένων

Στην περίπτωση των συμβάντων με διάρκεια μέσα σε ένα ταξίδι, η μορφή των δεδομένων που χρειάζεται ο αλγόριθμος είναι διαφορετική, καθώς και δυσκολότερη ως προς την “κατασκευή” της. Όπως και προηγουμένως, φιλτράρονται σε πρώτο στάδιο τα σημεία στα οποία πραγματοποιείται κάποιο συμβάν. Οι στήλες οι οποίες μένουν για επεξεργασία αφορούν το tripid, τον κωδικό του κάθε ταξιδιού που πραγματοποίησε ο χρήστης, καθώς και το osmid, που αφορά τον κωδικό ενός μέρους του δρόμου που εξετάζεται. Ο κωδικός αυτός περιέχει πολλά σημεία και συνεπώς διαφορετικές συντεταγμένες πάνω σε ένα σχετικά μικρό κομμάτι του δρόμου. Γίνεται ταξινόμηση ανά ταξίδι (tripid) και οι κωδικοί (osmids) που έχει κάθε ταξίδι συγχωνεύονται για κάθε ένα ξεχωριστό ταξίδι.

Τα δεδομένα, λοιπόν, που εισέρχονται στον αλγόριθμο αυτό έχουν την ακόλουθη μορφή:

Πίνακας 4. 5: Δεδομένα προς εισαγωγή στον αλγόριθμο frequent sequence mining.

tripid	osm_ids
0009E9F2-6EC8-4767-BD81-9A1D19C22335,	"[27512944.0, 27512944.0, 27512944.0, 27512944.0, 27512944.0, ..]"
000BB6E1-FF6C-44B9-8BB9-1F366808653D,	"[80510839.0, 80510839.0, 43782893.0, 43782893.0, 80510839.0, ..]"
0013A98D-0D16-4B7C-B253-9739CD1E6424,	"[249704497.0, 475874923.0, 475874923.0, 475874923.0, 475874923.0, 249704497.0, ..]"
00351DA7-008D-4A2B-A698-7D14D8F6557A,	"[nan, 435093165.0, 435093165.0, 306247268.0, 306247268.0, 437167075.0, ..]"

Η πρώτη στήλη αφορά τα ξεχωριστά ταξίδια που πραγματοποιούνται και η δεύτερη δείχνει όλα τα osmids που περιλαμβάνει κάθε ένα ταξίδι από αυτά μέσα στην αρχική βάση δεδομένων.

4.3.2 Ανάλυση Συμπεριφοράς

Όπως αναφέρθηκε και στη Μεθοδολογία που ακολουθήθηκε στην παρούσα εργασία, η μόνη παράμετρος που χρειάζεται να εισέλθει σε αλγορίθμους τύπου frequent sequence mining είναι ένας αριθμός, που εμφανίζεται ως η τιμή της μεταβλητής support. Πρόκειται για έναν αριθμό, ο οποίος υποδεικνύει τις ελάχιστες φορές όπου μπορεί ένα μοτίβο να εμφανίζεται στην αρχική βάση δεδομένων προκειμένου να θεωρηθεί ομάδα (pattern) στα αποτελέσματα του αλγορίθμου.

Για τους σκοπούς της παρούσας εργασίας, η τιμή της παραμέτρου αυτής προσδιορίστηκε δεδομένου του παρεχόμενου πλήθους δεδομένων στην αρχική βάση δεδομένων. Πραγματοποιήθηκαν, επίσης, αρκετές δοκιμές για τη βέλτιστη επιλογή της παραμέτρου. Έτσι, η παράμετρος αυτή λαμβάνει τη τιμή 30. Η τιμή αυτή είναι η ίδια και στην περίπτωση του ενός οδηγού και στην περίπτωση περισσότερων οδηγών.

Ο αλγόριθμος, ψάχνει να βρει ομάδες (patterns) από osmids, τα οποία εμφανίζονται συχνά μαζί μέσα στην αρχική βάση δεδομένων. Το αποτέλεσμα στο οποίο καταλήγει ο αλγόριθμος αποτελείται από αυτές τις ομάδες, συνοδευόμενες από έναν αριθμό (support) ,ο οποίος δείχνει το πόσες φορές εμφανίζονται στην αρχική βάση δεδομένων (κόκκινο χρώμα). Μέρος του αποτελέσματος είναι:

```
[ ((5196746.0,), 47), ((15371919.0,), 59), ((15371919.0, 15591903.0), 39),
((15371919.0, 23184757.0), 32), ((15371919.0, 30033771.0), 32), ((15371919.0,
30033771.0, 30033772.0), 31) ,... ]
```

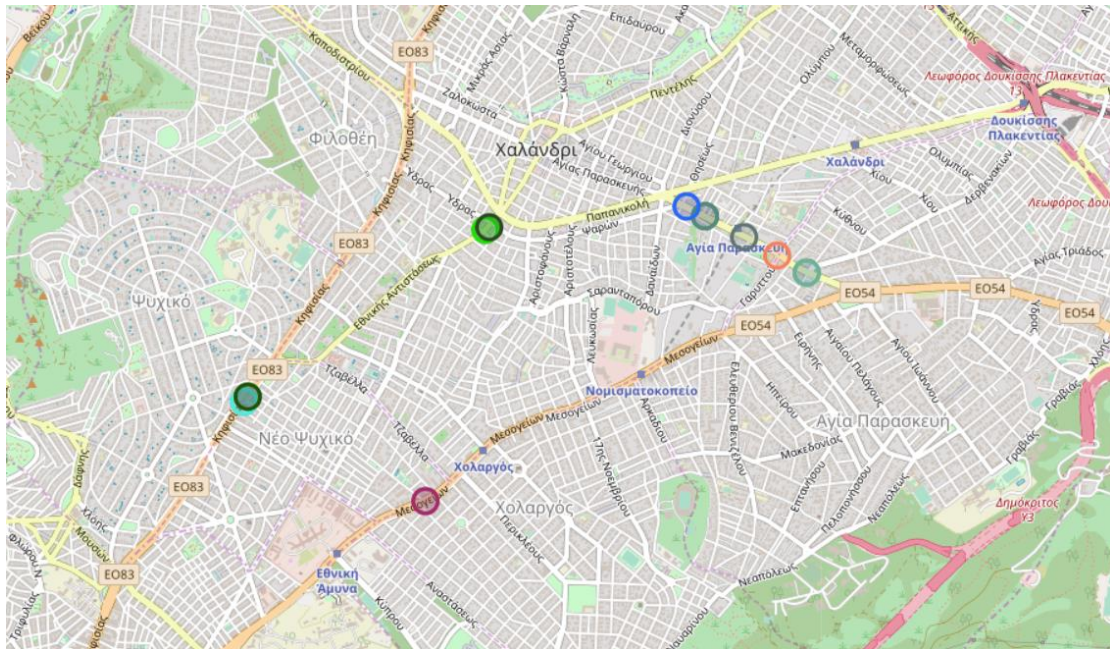
Κάθε μια από τις ομάδες που δημιουργούνται αποτελεί υποσύνολο κάποιας άλλης ομάδας.

Για την αξιολόγηση των ομάδων που δημιουργούνται γίνεται η αρχική απεικόνισή τους σε ένα χάρτη. Όπως παρατηρήθηκε, οι ομάδες που δημιουργούνται είναι πολλές και δεν γίνεται καμία διάκριση για τη σημαντικότητα αυτών πάνω στο οδικό δίκτυο.

Επιπλέον, μερικές από τις ομάδες πιθανόν να περιέχουν ορισμένα osmids, τα οποία απέχουν αρκετά μεταξύ τους (απόσταση μεγαλύτερη από 1.5 χιλιόμετρα). Έτσι, το αποτέλεσμα δεν μπορεί να αξιολογηθεί καταλλήλως. Για το λόγο αυτό, δημιουργείται αρχικά μια συνάρτηση για τον υπολογισμό της απόστασης των σημείων μέσα στην ίδια ομάδα, προκειμένου ο αλγόριθμος να θέτει προς επεξεργασία τα σημεία της ομάδας τα οποία βρίσκονται κοντά μεταξύ τους. Επιπλέον, προκειμένου να αξιολογηθούν τα πιο σημαντικά (επικίνδυνα) σημεία πάνω στο οδικό δίκτυο και να μειωθεί το μεγάλο πλήθος των ομάδων που δημιουργούνται, γίνεται μια ακόμα

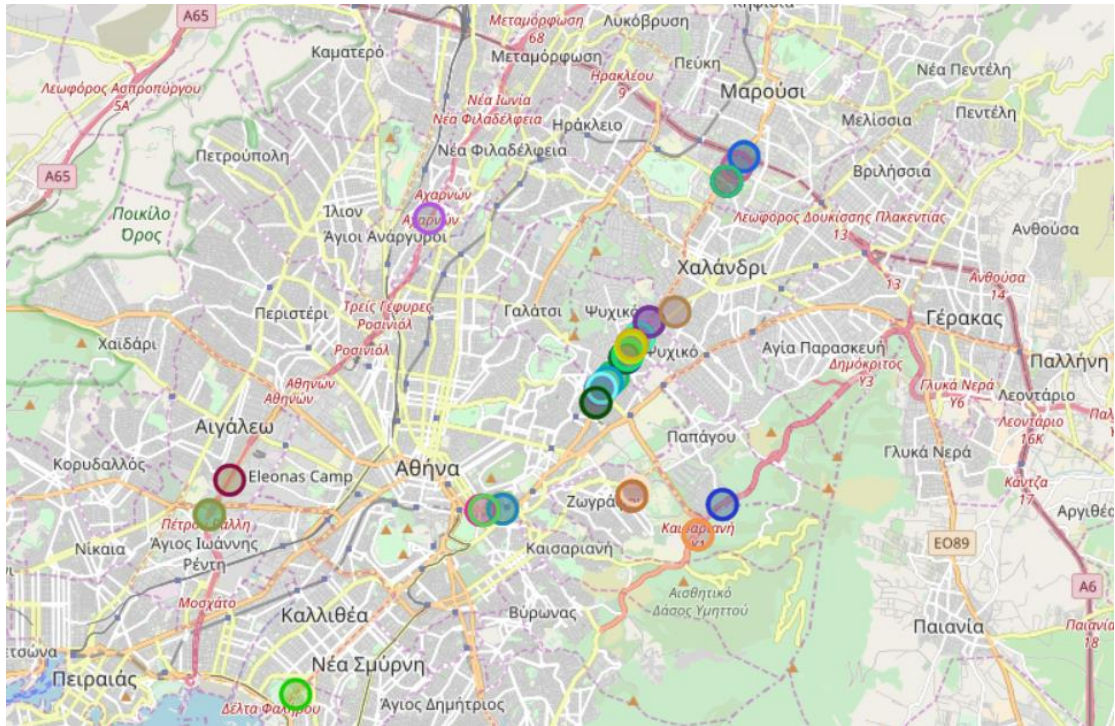
προσαρμογή του αλγορίθμου. Δημιουργείται μια συνάρτηση προκειμένου να βρεθεί η σημαντικότητα της κάθε ομάδας με σκοπό να αξιολογηθούν τα πιο επικίνδυνα σημεία του οδικού δικτύου. Η συνάρτηση αυτή, για κάθε μια από τις ομάδες που έχουν δημιουργηθεί, περιλαμβάνει το γινόμενο του πλήθους των osmids που περιέχει η ομάδα πολλαπλασιασμένο με τον αριθμό εμφάνισης της ομάδας αυτής. Για παράδειγμα, η ομάδα : ((15371919.0, 30033771.0), 32) αποκτά έναν αριθμό σημαντικότητας ίσο με δυο επί τριάντα δυο, δηλαδή εξήντα τέσσερα.

Εισάγοντας, λοιπόν, τις δυο αυτές επιπλέον παραμέτρους γίνεται η ταξινόμηση των ομάδων στο οδικό δίκτυο και προκύπτει η παρακάτω εικόνα για το σύνολο των διαδρομών σε επίπεδο ενός χρήστη και αντίστοιχα η επόμενη εικόνα σε επίπεδο διαφορετικών χρηστών.



Εικόνα 4. 12: Χάρτης επικινδυνότητας για συμβάντα υπέρβασης του ορίου ταχύτητας μεμονωμένου χρήστη.

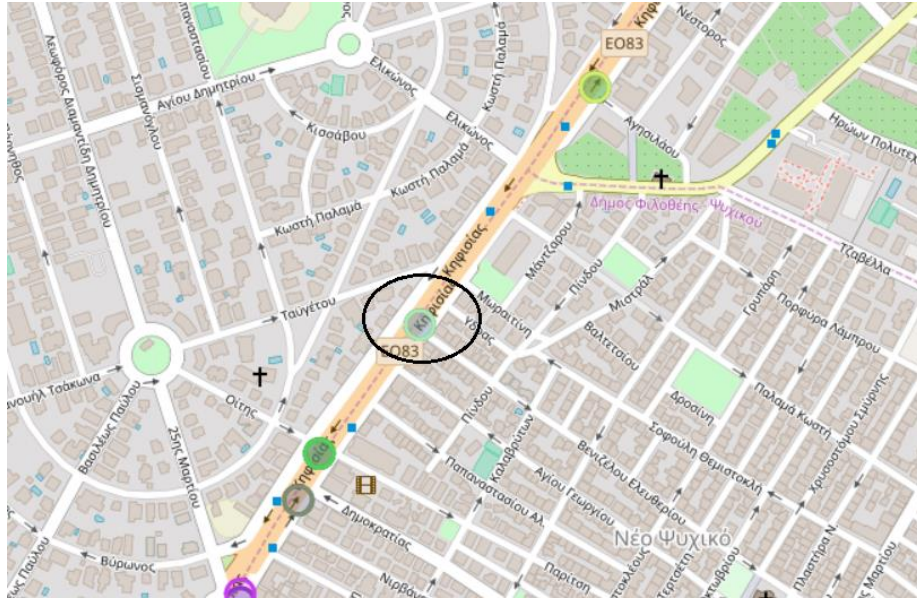
Στον παρακάτω χάρτη, απεικονίζονται οι ομάδες με τα πιο επικίνδυνα osmids που βρέθηκαν στην περίπτωση των 147 οδηγών. Κάθε διαφορετικό χρώμα αναπαριστά και μια διαφορετική ομάδα του οδικού δικτύου. Στο χάρτη φαίνονται οι ομάδες οι οποίες θεωρήθηκαν οι πιο σημαντικές από την εισαγωγή των δυο επιπλέον παραμέτρων που αναφέρθηκαν παραπάνω.



Εικόνα 4. 13: Χάρτης επικινδυνότητας για συμβάντα υπέρβασης του ορίου ταχύτητας πολλών χρηστών.

Από την χωρική αναπαράσταση των ομάδων που γίνονται συνολικά οι περισσότερες υπερβάσεις του ορίου ταχύτητας, παρατηρούνται τα εξής:

- Όπως ήταν αναμενόμενο, οι ομάδες που κρίθηκαν ως οι πιο επικίνδυνες βρίσκονται σε βασικές αρτηρίες του οδικού δικτύου της Αθήνας. Ο οδηγός είναι περισσότερο ευάλωτος στην ανάπτυξη μεγάλης ταχύτητας σε μεγαλύτερες αρτηρίες.
- Για τη βάση δεδομένων από 147 χρήστες, η Λεωφόρος Κηφισίας από το ύψος της οδού Εθνικής Αντιστάσεως έως και το ύψος της οδού Πανόρμου παρουσιάζεται ως ένα από τα πιο επικίνδυνα τμήματα για την υπέρβαση του ορίου ταχύτητας. Είναι χρήσιμο να διερευνηθούν οι παράγοντες που οδηγούν στην ανάπτυξη της συμπεριφοράς αυτής από τους οδηγούς.
- Συγκρίνοντας τις απεικονίσεις των δυο διαφορετικών βάσεων δεδομένων παρατηρείται μια ομοιομορφία ως προς τη συμπεριφορά των οδηγών. Πιο συγκεκριμένα, η Λεωφόρος Κηφισίας στο ύψος της Φιλοθέης αποτελεί τμήμα του δρόμου όπου γίνεται συχνή υπέρβαση ταχύτητας. Αυτό δεν αποτελεί τυχαίο γεγονός, καθώς φαίνεται και από τη συμπεριφορά του ενός χρήστη αλλά και από τη συμπεριφορά διαφορετικών χρηστών. Η ομάδα αυτή φαίνεται στο παρακάτω χάρτη:



Εικόνα 4. 14: Κρίσιμη ομάδα για συμβάντα υπέρβασης του ορίου ταχύτητας.

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

5.1 ΕΙΣΑΓΩΓΗ

Η εισαγωγή νέων τεχνολογιών καθιστά δυνατή τη συλλογή, αποθήκευση και μετάδοση δεδομένων μεγάλης ακρίβειας μέσω έξυπνων κινητών τηλεφώνων. Η βιβλιογραφία κατέστη επαρκής όσον αφορά στην αξιοποίηση αυτών των δεδομένων για την ομαδοποίηση συμπεριφορών ή γενικότερα δεδομένων. Ωστόσο, δεν υπήρχαν επαρκείς αναφορές για τη χωρική αναπαράσταση αυτών. Έτσι, επιδιώχθηκε η δημιουργία μιας μεθοδολογικής προσέγγισης που αρχικά ομαδοποιεί τα δεδομένα βάσει συμπεριφορών και έπειτα πραγματοποιεί τη χωρική απεικόνιση αυτών σε χάρτες επικινδυνότητας.

Η μεθοδολογία που ακολουθήθηκε για προσέγγιση του ζητήματος αφορά την ανάπτυξη αλγορίθμων ομαδοποίησης διαφορετικών για σημειακά συμβάντα και για συμβάντα τα οποία έχουν διάρκεια μέσα σε ένα ταξίδι. Μέσω της διαδικασίας αυτή, δημιουργούνται ομάδες που αποτελούνται από σημεία στα οποία οι οδηγοί αντέδρασαν λανθασμένα κατά τη διάρκεια της οδήγησης, παρατηρήθηκε δηλαδή ένα συμβάν (event).

Έπειτα, πραγματοποιήθηκε η χωρική απεικόνιση των ομάδων αυτών σε χάρτες επικινδυνότητας, οι οποίοι αποτελούν εργαλείο ενημέρωσης των χρηστών αλλά και του συστήματος διαχείρισης της κυκλοφορίας.

5.2 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Πρωταρχικό συμπέρασμα αποτελεί η πραγματοποίηση του σκοπού της παρούσας διπλωματικής εργασίας. Κριτήριο για την αξιολόγηση της επιτυχίας της διαδικασίας αποτελούν οι χάρτες επικινδυνότητας και η εγκυρότητα των επικίνδυνων σημείων του οδικού δικτύου πάνω σε αυτούς. Η παροχή των δεδομένων είναι σκόπιμο να γίνεται εφαρμόζοντας νέες τεχνολογίες έξυπνων κινητών τηλεφώνων. Η χρήση αυτών είναι πλέον διαδεδομένη, καθώς και εύκολη για κάθε άτομο.

Κατ' αρχήν υλοποιήθηκε αλγόριθμος ομαδοποίησης των συμπεριφορών διαφόρων οδηγών εντός του οδικού δικτύου της Αθήνας και σε επόμενη φάση έγινε η απεικόνιση των ομάδων που προέκυψαν σε χάρτες επικινδυνότητας του δικτύου. Οι χάρτες αυτοί αποτελούν εργαλείο για τους οδηγούς ή για ένα σύστημα διαχείρισης της κυκλοφορίας για τους λόγους που παρουσιάζονται παρακάτω.

Οι χάρτες επικινδυνότητας αποτελούν εργαλείο πληροφόρησης και προειδοποίησης για τον κάθε χρήστη ξεχωριστά. Μέσω αυτών, ο οδηγός μπορεί να είναι ενήμερος για τα τμήματα του δικτύου όπου πραγματοποιούνται τα περισσότερα οδηγικά συμβάντα, προκειμένου να λαμβάνει εγκαίρως τις σωστές αποφάσεις.

Οι έρευνες συγκοινωνιακών ζητημάτων, μέσω των χαρτών επικινδυνότητας, ενισχύονται. Γίνονται γνωστά τα οδικά τμήματα του δικτύου όπου συμβαίνουν τα περισσότερα συμβάντα και έτσι γίνεται ευκολότερη η αξιολόγηση των λόγων για τους οποίους συμβαίνουν στις θέσεις αυτές (για παράδειγμα, λόγω ανεπαρκούς σηματοδότησης ή λανθασμένης γεωμετρίας της οδού).

Οι χάρτες επικινδυνότητας αποτελούν εργαλείο διαχείρισης της κυκλοφορίας του οδικού δικτύου.

Οι αλγόριθμοι που αναπτύχθηκαν είναι δυνατό να προσαρμοστούν σε διαφορετικά σημεία, δεν περιορίζονται δηλαδή για το οδικό δίκτυο της Αθήνας. Το ίδιο ισχύει και για τους χάρτες επικινδυνότητας.

Μέσω της απεικόνισης της συμπεριφοράς του οδηγού στο δίκτυο, ο ίδιος μέσα από την εικόνα είναι εύκολο να κατανοήσει τα σημεία στα οποία πραγματοποίησε λάθος και να ξέρει να ενεργήσει καταλληλότερα την επόμενη φορά που θα περάσει από το ίδιο σημείο του δικτύου.

Οι χάρτες αποτελούν εργαλείο αξιοποιήσιμο για οργανισμούς που συμβάλλουν στη βελτίωση της οδηγικής συμπεριφοράς.

Σχετικά με τη συμπεριφορά των οδηγών εντός του οδικού δικτύου της Αθήνας, παρατηρείται ότι:

- Η συμπεριφορά των οδηγών επηρεάζεται σημαντικά από την ύπαρξη φωτεινής σηματοδότησης. Όπως φαίνεται και από τους κρίσιμους κόμβους του Κεφαλαίου 4, οι οδηγοί κοντά στη διασταύρωση οδών αντιδρούν επιθετικά, είτε επιταχύνουν είτε επιβραδύνουν αναλόγως την εκάστοτε σηματοδότηση.
- Οι ταχύτητες στις οποίες πραγματοποιείται το μεγαλύτερο ποσοστό των συμβάντων δεν είναι τόσο σημαντικές, προκειμένου να θεωρηθούν επικίνδυνες και να προκαλέσουν τυχόν ατύχημα.
- Οι συμπεριφορές των οδηγών σχετίζονται μεταξύ τους, καθώς παρατηρήθηκε ομοιομορφία στις αντιδράσεις τους.

5.3 ΠΡΟΕΚΤΑΣΕΙΣ ΈΡΕΥΝΑΣ

Στην παρούσα διπλωματική εργασία η επίλυση του προβλήματος βασίστηκε στην καταγραφή αλγορίθμων για τον έλεγχο της ομαδοποίησης μεταξύ διαφορετικών σημείων του δικτύου και στην απεικόνιση αυτών σε χάρτες επικινδυνότητας. Κατά τη διάρκεια της εργασίας έγιναν ορισμένες προσεγγίσεις και παραδοχές για σαφή ορισμό του προβλήματος. Λόγω των περιορισμών που τέθηκαν αλλά και των προβληματισμών που προέκυψαν κατά τη διάρκεια της έρευνας κρίνεται απαραίτητη η παρουσίαση των σημείων όπου χρειάζεται περαιτέρω εμβάθυνση.

Πρωτίστως, όσον αφορά στα συμβάντα τα οποία δεν έχουν διάρκεια, σημειακά δηλαδή συμβάντα, όπως η απότομη επιτάχυνση, είναι σκόπιμο να πραγματοποιηθεί περαιτέρω ανάλυση σε δεύτερο στάδιο. Δηλαδή, αφού έχει γίνει η ομαδοποίηση των ακραίων οδηγικών συμβάντων και έχουν κατανεμηθεί σε ομάδες, θα πραγματοποιηθεί μια ακόμα ομαδοποίηση των σημείων αυτών που έχουν καταταγεί σε κάποια ομάδα. Με τον τρόπο αυτό, γίνεται περαιτέρω ανάλυση και έτσι ο ερευνητής είναι σε θέση να αναγνωρίσει σημεία του οδικού δικτύου τα οποία έχουν ακόμα μεγαλύτερο βαθμό επικινδυνότητας από τα προηγούμενα.

Για οδηγικά συμβάντα, τα οποία έχουν διάρκεια, όπως η υπέρβαση του ορίου ταχύτητας σε ένα τμήμα του οδικού δικτύου, πρέπει να ληφθεί υπόψιν και να αξιολογηθεί με τον κατάλληλο τρόπο ο παράγοντας του χρόνου. Οι (Birant & Kut, 2007) έχουν κατασκευάσει έναν αλγόριθμο ικανό να πραγματοποιήσει το παραπάνω. Πρόκειται για τον αλγόριθμο STDBSCAN. Είναι ένας αλγόριθμος ομαδοποίησης των δεδομένων, εξέλιξη του αλγορίθμου DBSCAN, ο οποίος έχει τη δυνατότητα ομαδοποίησης όχι μόνο χωρικών δεδομένων αλλά και χωροχρονικών. Με τον τρόπο αυτό, θα γίνει αντιληπτή η χωροχρονική απεικόνιση των συμβάντων αυτών. Επιπλέον, αποτελεί και εργαλείο αξιολόγησης της ήδη υπάρχουσας επίλυσης που έγινε στην παρούσα διπλωματική εργασία (αλγόριθμος τύπου frequent sequence mining). Τα αποτελέσματα των δύο αυτών επιλύσεων μπορούν να συγκριθούν και να γίνει παρουσίαση των κατάλληλων συμπερασμάτων.

Σύμφωνα με την προηγούμενη προσέγγιση, η (Moraga, 2017) έχει κατασκευάσει μια διαδικτυακή εφαρμογή σε γλώσσα προγραμματισμού R μέσω της οποίας ανιχνεύονται και ταυτόχρονα απεικονίζονται σε χάρτη ομάδες χωροχρονικών δεδομένων που παρουσιάζουν ομοιομορφία στη συμπεριφορά. Τα αποτελέσματα της προσέγγισης αυτής αφορούν χάρτες επιδημιολογίας στους οποίους φαίνεται η αντίστοιχη επικινδυνότητα κάθε τμήματος ανάλογα την περιοχή που εξετάζεται. Στα πλαίσια της παρούσας διπλωματικής εργασίας, η υλοποίηση των δύο παραπάνω αλγορίθμων δεν ήταν εφικτή λόγω έλλειψης όγκου δεδομένων και μετέπειτα επεξεργασία αυτών.

Επιπλέον, σημαντικός παράγοντας που δεν πρέπει να παραλειφθεί σε επόμενη σχετική έρευνα που βασίζεται στην οδηγική συμπεριφορά των οδηγών αποτελεί η γεωμετρία της οδού. Στα πλαίσια της παρούσας διπλωματικής ο παράγοντας αυτός δεν λήφθηκε υπόψιν για την ανάλυση της οδηγικής συμπεριφοράς των οδηγών. Ωστόσο, η γεωμετρία της οδού αποτελεί καθοριστικό παράγοντα κατά τη διάρκεια της οδήγησης. Είναι σημαντική για την έγκαιρη και ορθή λήψη αποφάσεων από τους οδηγούς.

Επίσης σημαντικός παράγοντας που επηρεάζει την οδηγική συμπεριφορά είναι ο κυκλοφοριακός φόρτος κατά τη διάρκεια ενός ταξιδιού. Ο κυκλοφοριακός φόρτος αποτελεί συνάρτηση πολλών παραγόντων και είναι δύσκολο να συμπεριληφθεί σε τέτοιου είδους έρευνες. Χρειάζονται επιπλέον στοιχεία οδηγικών συμπεριφορών είτε για ώρες αιχμής είτε όχι. Επιπρόσθετα, η συμπεριφορά των οδηγών επηρεάζεται και από τις κινήσεις των λοιπών οδηγών που κινούνται στο ίδιο τμήμα της οδού. Είναι ένας ακόμη παράγοντας που είναι χρήσιμο να ληφθεί υπόψιν για αξιολόγηση και διερεύνηση της οδηγικής συμπεριφοράς.

Τέλος, οι καιρικές συνθήκες της εκάστοτε περιοχής αποτελούν σημαντικό παράγοντα για τη λήψη αποφάσεων κατά τη διάρκεια της οδήγησης. Επομένως, κρίνεται χρήσιμη η εισαγωγή του παράγοντα αυτού σε περαιτέρω ανάλυση και η διερεύνηση των επιπτώσεων που προκαλεί.

6. ΒΙΒΛΙΟΓΡΑΦΙΑ

- Assum, T. (1997). Attitudes and road accident risk. *Accident Analysis & Prevention*, 29(2), 153–159. [https://doi.org/10.1016/s0001-4575\(96\)00071-1](https://doi.org/10.1016/s0001-4575(96)00071-1)
- Birant, D., & Kut, A. (2007). ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data and Knowledge Engineering*, 60(1), 208–221. <https://doi.org/10.1016/j.datak.2006.01.013>
- Bollen, K. A. (1989). Structural equations with latent variables. *Wiley Series in Probability and Mathematical Statistics*, 8, 528. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., & Ratti, C. (2011). Real-time urban monitoring using cell phones: A case study in Rome. *IEEE Transactions on Intelligent Transportation Systems*, 12(1), 141–151. <https://doi.org/10.1109/TITS.2010.2074196>
- Chowdhury, A., Chakravarty, T., & Balamuralidhar, P. (2014). A novel approach to improve vehicle speed estimation using smartphone's INS/GPS sensors. In *Proceedings of the International Conference on Sensing Technology, ICST* (Vol. 2014–Janua).
- Dang, V. C., Kubo, M., Sato, H., & Namatame, A. (2015). Building Safety Road Maps Based on Difference of Judgment of Road Users by their Smartphone. *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, 6(9), 15–23.
- Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring Social Network Structure using Mobile Phone Data. *Pnas*, 106(usually 1), 15274–15278. <https://doi.org/10.1073/pnas.0900282106>
- Estelles-Arolas, E., & Gonzalez-Ladron-De-Guevara, F. (2012). Tasks-based classification of crowdsourcing initiatives. *PROFESIONAL DE LA INFORMACION*, 21(3), 283–291.
- Fournier-viger, P., & Lin, J. C. (2017). Survey of sequential pattern mining. *Data Science and Pattern Recognition*, 1(1), 0–4.

- Garling, T., Gillholm, R., & Garling, A. (1998). Reintroducing attitude theory in travel behavior research. *Transportation*, 25, 129–146.
- Han, Q., Zhu, Y., Zeng, L., Ye, L., He, X., Liu, X., ... Zhu, Q. (2015). A Road Hotspots Identification Method Based on Natural Nearest Neighbor Clustering. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* (Vol. 2015–Octob, pp. 553–557). <https://doi.org/10.1109/ITSC.2015.97>
- Handel, P., Skog, I., Wahlstrom, J., Bonawiede, F., Welch, R., Ohlsson, J., & Ohlsson, M. (2014). Insurance telematics: Opportunities and challenges with the smartphone solution. *IEEE Intelligent Transportation Systems Magazine*, 6(4), 57–70. <https://doi.org/10.1109/MITS.2014.2343262>
- Hoffmann, E. R., & Mortimer, R. G. (1994). Drivers' estimates of time to collision. *Accident Analysis and Prevention*, 26(4), 511–520. [https://doi.org/10.1016/0001-4575\(94\)90042-6](https://doi.org/10.1016/0001-4575(94)90042-6)
- Kang, C., Ma, X., Tong, D., & Liu, Y. (2012). Intra-urban human mobility patterns: An urban morphology perspective. *Physica A: Statistical Mechanics and Its Applications*, 391(4), 1702–1717. <https://doi.org/10.1016/j.physa.2011.11.005>
- Kiefer, R. J., Flannagan, C. A., & Jerome, C. J. (2006). Time-to-Collision Judgments Under Realistic Driving Conditions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 48(2), 334–345. <https://doi.org/10.1518/001872006777724499>
- Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., & Campbell, A. T. (2010). A survey of mobile phone sensing. *IEEE Communications Magazine*, 48(9), 140–150. <https://doi.org/10.1109/MCOM.2010.5560598>
- Lazer, D., Pentland, A., Adamic, L. A., Aral, S., Barabási, A. L., Brewer, D., ... Alstynne, M. (2009). Computational Social Science. *Science*, 323, 721–723. <https://doi.org/10.1126/science.1167742>
- Liu, Y. ., Kang, C. ., & Wang, F. . (2014). Towards big data-driven human mobility patterns and models. *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*, 39(6), 660–666. <https://doi.org/10.13203/j.whugis20140149>

- Minderhoud, M. M., & Bovy, P. H. L. (2001). Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis and Prevention*, 33(1), 89–97. [https://doi.org/10.1016/S0001-4575\(00\)00019-1](https://doi.org/10.1016/S0001-4575(00)00019-1)
- Moraga, P. (2017). SpatialEpiApp: A Shiny web application for the analysis of spatial and spatio-temporal disease data. *Spatial and Spatio-Temporal Epidemiology*, 23, 47–57. <https://doi.org/10.1016/j.sste.2017.08.001>
- Paefgen, J., Kehr, F., Zhai, Y., & Michahelles, F. (2012). Driving behavior analysis with smartphones: insights from a controlled field study. *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, (January), 36:1–36:8. <https://doi.org/10.1145/2406367.2406412>
- Ranjith, R., Athanesious, J. J., & Vaidehi, V. (2016). Anomaly detection using DBSCAN clustering technique for traffic video surveillance. In *ICoAC 2015 - 7th International Conference on Advanced Computing*. <https://doi.org/10.1109/ICoAC.2015.7562795>
- Semiring, R. W., & Zain, J. M. (2010). Cluster Evaluation of Density Based Subspace Clustering. *Journal of Computing*, 2(11), 2151–9617.
- Soper, D. (2012). Viewpoint: Is human mobility tracking a good idea? In *Communications of the ACM* (Vol. 55, pp. 35–37). <https://doi.org/10.1145/2133806.2133819>
- Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. *Transportation Research Part C: Emerging Technologies*, 16(3), 320–331. <https://doi.org/10.1016/j.trc.2008.01.001>
- Vlahogianni, E. I., & Barmounakis, E. N. (2017). Driving analytics using smartphones: Algorithms, comparisons and challenges. *Transportation Research Part C: Emerging Technologies*, 79, 196–206. <https://doi.org/10.1016/j.trc.2017.03.014>
- Weekley, J., Barrell, J., & McCarthy, T. (2016). Developing a Road Safety Review Tool to Identify Design Standard and Safety Deficits on High Risk Road Sections. In *Transportation Research Procedia* (Vol. 14, pp. 4130–4139). <https://doi.org/10.1016/j.trpro.2016.05.384>
- Zhao, Y. (2000). Mobile Phone Location Determination and Its Impact on Intelligent Transportation Systems. *IEEE Transactions on Intelligent Transportation Systems*, 1(1), 55–64. <https://doi.org/10.1109/6979.869021>

- Zheng, Y., Liu, Y., Yuan, J., & Xie, X. (2011). Urban computing with taxicabs. In *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11* (p. 89). <https://doi.org/10.1145/2030112.2030126>
- Zou, X., & Zhu, Q. (2011). Abnormal structure in regular data revealed by Isomap with natural nearest neighbor. In *Communications in Computer and Information Science* (Vol. 216 CCIS, pp. 538–544). https://doi.org/10.1007/978-3-642-23345-6_97