



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

ΑΛΓΟΡΙΘΜΟΙ ΕΝΙΣΧΥΤΙΚΗΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗ ΒΕΛΤΙΩΣΗ ΤΗΣ ΟΔΗΓΙΚΗΣ ΣΥΜΠΕΡΙΦΟΡΑΣ

Διπλωματική Εργασία

ΒΑΣΙΛΗΣ ΚΥΡΙΑΖΟΠΟΥΛΟΣ

Επιβλέπουσα: Ελένη Ι. Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια
Ε.Μ.Π.

Αθήνα, Οκτώβριος 2019

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της Διπλωματικής μου Εργασίας, θα ήθελα να ευχαριστήσω όσους με υποστήριξαν και με βοήθησαν να κλείσω αυτόν τον σημαντικό κύκλο της ζωής μου.

Αρχικά, θα ήθελα να ευχαριστήσω θερμά την κ. Ελένη Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια της Σχολής Πολιτικών Μηχανικών Ε.Μ.Π., για την δυνατότητα που μου προσέφερε να εκπονήσω ένα τόσο ενδιαφέρον θέμα και για την υποστήριξή της σε όλα τα στάδια υλοποίησής του. Εκτιμάω την πολύτιμη καθοδήγηση και βοήθεια που μου προσέφερε, καθώς και την εξαιρετική συνεργασία μας.

Θα ήθελα επίσης να ευχαριστήσω τον Διδάκτορα Εμμανουήλ Μπαρμπουνάκη ο οποίος μου έδωσε τα δεδομένα για την υλοποίηση της Εργασίας και ήταν διαθέσιμος να μου προσφέρει οδηγίες και καθοδήγηση όποτε ήταν αναγκαίο.

Τέλος θα ήθελα να ευχαριστήσω όλους τους κοντινούς μου ανθρώπους που ήταν δίπλα μου και με στήριξαν σε κάθε βήμα των σπουδών μου και κυρίως την οικογένεια μου.

Αθήνα, Οκτώβριος 2019

Βασίλης Κυριαζόπουλος

ΑΛΓΟΡΙΘΜΟΙ ΕΝΙΣΧΥΤΙΚΗΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗ ΒΕΛΤΙΩΣΗ ΤΗΣ ΟΔΗΓΙΚΗΣ ΣΥΜΠΕΡΙΦΟΡΑΣ

Βασίλης Κυριαζόπουλος

Επιβλέπουσα: Ελένη Ι. Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια
Ε.Μ.Π.

ΣΥΝΟΨΗ

Τα τελευταία χρόνια η εξέλιξη της τεχνολογίας έχει επιτρέψει την εισαγωγή στα οχήματα συστημάτων αυτοματισμού τα οποία βελτιώνουν την οδική ασφάλεια μειώνοντας την συμβολή του ανθρώπινου παράγοντα στην διαδικασία της οδήγησης. Στόχος της παρούσας Διπλωματικής Εργασίας είναι η δημιουργία ενός αλγορίθμου ενισχυτικής μάθησης, που θα αξιολογεί το επίπεδο της οδικής ασφάλειας και θα προτείνει στον οδηγό ενέργειες για την αποφυγή συγκρούσεων, με σκοπό τη βελτίωση της οδηγικής συμπεριφοράς. Τα δεδομένα για την εκπαίδευση του αλγορίθμου λήφθηκαν από το πείραμα που πραγματοποιήθηκε σε κεντρικούς οδικούς άξονες της Αθήνας, όπου καταγράφηκε η κυκλοφορία με χρήση ΣμηΕΑ και εξήχθησαν οι τροχιές των οχημάτων. Για την διαμόρφωση του τελικού αλγορίθμου Q-Μάθησης έγιναν δοκιμές για διάφορους συνδυασμούς του ποσοστού εξερεύνησης, του εκπαιδευτικού παράγοντα γ και του βαθμού εκμάθησης. Ο τελικός Πίνακας-Q που διαμορφώθηκε είναι σε θέση να χρησιμοποιηθεί εντός οχημάτων για τη βελτίωση της οδικής ασφάλειας και την μείωση των συγκρούσεων. Η περαιτέρω αξιολόγηση της ποιότητας των αποτελεσμάτων κατέδειξε την ανάγκη για εκπαίδευση σε μεγαλύτερο δείγμα, ώστε να εξαλειφθούν κάποιες καταστάσεις που ο αλγόριθμος δεν προτείνει τη βέλτιστη πολιτική οδήγησης.

Λέξεις κλειδιά: οδική ασφάλεια, ενισχυτική μάθηση, προηγμένα συστήματα υποβοήθησης οδηγού, Q-Μάθηση, αποφυγή συγκρούσεων, ακολουθούντα οχήματα, συμπεριφορά οδηγού

REINFORCEMENT LEARNING ALGORITHMS FOR IMPROVING DRIVING BEHAVIOR

Vasilis Kyriazopoulos

Supervisor: Eleni I. Vlahogianni, Associate Professor NTUA

ABSTRACT

In recent years, the evolution of technology has allowed the introduction of automation systems in vehicles, that improve road safety by reducing the contribution of the human factor to the driving process. The objective of this Diploma Thesis is to propose a reinforcement learning algorithm for evaluating the level of safety of driving in relation to the driving environment and proposing actions to the driver in order to avoid a collision from occurring and to improve driving behavior. The data for the training of the algorithm were sourced from the study that was conducted in an extended region of the Athens downtown area, where the traffic was recorded using a swarm of UAVs and the vehicle trajectories were extracted. A Q-learning algorithm is developed and various combinations of the exploration rate, the discount factor γ and the learning rate were tested for the optimal algorithm. The final Q-Table can be used inside vehicles for collision avoidance in order to improve road safety. Results indicate that apart from the exceptional results, more training is needed to tackle some cases where the algorithm provided suboptimal actions that did not improve road safety.

Keywords: road safety, reinforcement learning, advanced driver assistance systems, Q-Learning, collision avoidance, car following, driver behavior

ΠΕΡΙΛΗΨΗ

Η ραγδαία ανάπτυξη της τεχνολογίας των μη επανδρωμένων αεροσκαφών τα τελευταία χρόνια έχει οδηγήσει σε σημαντικές προόδους στη συλλογή και επεξεργασία κυκλοφοριακών δεδομένων. Οι τροχιές των οχημάτων που προκύπτουν μπορούν να χρησιμοποιηθούν για την μικροσκοπική ανάλυση διαφόρων κυκλοφοριακών φαινομένων και την εξαγωγή συμπερασμάτων σχετικά με την οδική ασφάλεια. Στόχος της παρούσας Διπλωματικής Εργασίας είναι η δημιουργία ενός αλγορίθμου ενισχυτικής μάθησης για τη βελτίωση της οδηγικής συμπεριφοράς. Συγκεκριμένα, ένας πράκτορας αλληλοεπιδρά με το περιβάλλον της οδήγησης και κάθε χρονική στιγμή αναγνωρίζει το επίπεδο της οδικής ασφάλειας και προτείνει ενέργειες για την αποφυγή συγκρούσεων και τη βελτίωση της οδικής ασφάλειας.

Στο πλαίσιο της βιβλιογραφικής ανασκόπησης αναδείχθηκαν τα πλεονεκτήματα και οι περιορισμοί της επιλεχθείσας μεθοδολογίας, καθώς και οι ελλείψεις των ερευνών με σχετικό αντικείμενο, έτσι ώστε να οριστεί πλήρως το τελικό θέμα. Από την έρευνα των συναφών εργασιών προέκυψε ότι η ενισχυτική μάθηση χρησιμοποιείται σε πολλά συστήματα αυτόματης οδήγησης καθώς είναι σε θέση να διαχειριστεί πολυδιάστατα περιβάλλοντα τα οποία δύσκολα περιγράφονται από καθορισμένο αριθμό καταστάσεων. Ακόμη, οι αλγόριθμοι που απασκοπούν στη βελτίωση της οδικής ασφάλειας πολλές φορές δεν εκπαιδεύονται με πραγματικά δεδομένα και δεν λαμβάνουν υπόψιν τα πλευρικά οχήματα, τα οποία μπορεί να επηρεάσουν την συμπεριφορά του οδηγού.

Η συλλογή των δεδομένων πραγματοποιήθηκε σε κεντρικούς άξονες της Αθήνας στο πλαίσιο μιας έρευνας για την συλλογή δεδομένων σε ένα εκτεταμένο δίκτυο προκειμένου να αναδειχθούν τα ιδιαίτερα χαρακτηριστικά των ΣμηΕΑ και η δυνατότητα να ξεπεράσουν υφιστάμενους περιορισμούς στην παρακολούθηση και καταγραφή κυκλοφοριακών ροών. Κατά την διάρκεια του πειράματος, 10 ΣμηΕΑ τοποθετήθηκαν σε καθορισμένες θέσεις και κατέγραφαν την κυκλοφορία. Από τα βίντεο εξήχθησαν οι τροχιές των οχημάτων καθώς και διάφορα κυκλοφοριακά μεγέθη, τα οποία στην συνέχεια τροποποιήθηκαν για την εισαγωγή στον τελικό αλγόριθμο. Από τις συντεταγμένες, οι οποίες ήταν στο σύστημα WGS84, υπολογίστηκαν οι αποστάσεις μεταξύ του εξεταζόμενου και των κοντινότερων γύρω οχημάτων, καθώς με βάση τις αποστάσεις και τις ταχύτητες της γύρω κυκλοφορίας καθορίζεται η κατάσταση στην οποία βρίσκεται το οδικό περιβάλλον. Συνολικά χρησιμοποιήθηκαν 220000 παραδείγματα για εκπαίδευση και 54976 για έλεγχο των αποτελεσμάτων.

Το περιβάλλον περιγράφεται από 72 καταστάσεις με βάση τον χρόνο για σύγκρουση (Time to Collision, TTC) του εξεταζόμενου με το προπορευόμενο όχημα, τον χρόνο για σύγκρουση του ακολουθούντος με το εξεταζόμενο όχημα και τις αποστάσεις των οχημάτων δεξιά και αριστερά του εξεταζόμενου. Ο χρόνος για σύγκρουση λαμβάνει υπόψιν και τον χρόνο αντίδρασης του οδηγού και θεωρείται κρίσιμος για τιμές μικρότερες των 3.9 δευτερολέπτων. Οι πλευρικές αποστάσεις εξετάζονται για κάθε συνδυασμό οχημάτων και θεωρούνται κρίσιμες όταν δεν υπερβαίνουν το 84.14% της κατανομής όλων των πλευρικών αποστάσεων του κάθε συνδυασμού οχημάτων. Οι διαθέσιμες ενέργειες είναι όλοι οι συνδυασμοί ενεργειών που μπορεί να κάνει ο οδηγός και διακρίνονται σε δύο επίπεδα επιτάχυνση/τίποτα/επιβράδυνση και αριστερά/τίποτα/δεξιά.

Για την διαμόρφωση του τελικού αλγορίθμου Q-Μάθησης πραγματοποιήθηκαν δοκιμές για διάφορους συνδυασμούς παραμέτρων. Στον τελικό αλγόριθμο, ο εκπαιδευτικός παράγοντας είναι ίσος με 0.9, ο ρυθμός εκμάθησης είναι ίσος με 0.2 και ο βαθμός εξερεύνησης ίσος με 0.2. Με βάση αυτές τις παραμέτρους προέκυψε ο τελικός Πίνακας-Q ο οποίος παρουσιάζεται στην συνέχεια. Κατά τον έλεγχο ο αλγόριθμος έκανε 4 λάθη, αρκετά μικρός αριθμός πράγμα που σημαίνει ότι ο αλγόριθμος είναι σε θέση να ανταπεξέλθει σε πραγματικές συνθήκες και να χρησιμοποιηθεί εντός οχημάτων για τη βελτίωση της οδηγικής συμπεριφοράς.

Από τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας προέκυψαν αποτελέσματα άμεσα συνδεδεμένα με τον κύριο στόχο που είχε τεθεί αρχικά. Τα σημαντικότερα συμπεράσματα που προκύπτουν είναι τα εξής:

Ο αλγόριθμος έχει την ικανότητα να μαθαίνει γρήγορα όπως προκύπτει και από το διάγραμμα της αξίας της μέσης μέγιστης ενέργειας όλων των καταστάσεων το οποίο συγκλίνει. Ένας αλγόριθμος ενισχυτικής μάθησης, για να γίνει αποδεκτός και να χρησιμοποιηθεί στην πράξη, θα πρέπει να φτάσει σε σύγκλιση κατά την διάρκεια της εκπαίδευσης. Η ταχύτητα με την οποία φτάνει στην σύγκλιση αποτελεί ένδειξη της ποιότητάς του. Συνεπώς ο αλγόριθμος που αναπτύχθηκε στα πλαίσια αυτής της Διπλωματικής Εργασίας μπορεί να θεωρηθεί ικανοποιητικός καθώς συγκλίνει αρκετά γρήγορα.

Όλοι οι αλγόριθμοι ενισχυτικής μάθησης έχουν την ικανότητα να επεκταθούν και να συνεχίσουν να μαθαίνουν χωρίς να υπάρχει ανάγκη δημιουργίας ενός νέου αλγορίθμου από την αρχή σε αντίθεση με τα στατιστικά μοντέλα. Έτσι λοιπόν, **ο αλγόριθμος που αναπτύχθηκε μπορεί να εκπαιδευτεί περαιτέρω** για να είναι σε θέση να αντιμετωπίζει με καλύτερο τρόπο τις κρίσιμες καταστάσεις. Οι αλγόριθμοι μαθαίνουν ακόμα και κατά την διάρκεια της χρήσης τους αφού ολοκληρωθεί η εκπαίδευση καθώς εκτίθενται σε καινούργιες καταστάσεις, βελτιώνοντας έτσι τις επιδόσεις τους και την εμπειρία.

Η συγκεκριμένη Διπλωματική Εργασία αποτελεί μια από τις πρώτες έρευνες που **λαμβάνουν υπόψιν και τα πλευρικά οχήματα** κατά την διάρκεια της αξιολόγησης του επιπέδου ασφάλειας της οδικής κατάστασης. Οι περισσότερες μελέτες πάνω στο αντικείμενο των συστημάτων αποφυγής σύγκρουσης εστιάζουν στο μπροστά και στο πίσω όχημα αγνοώντας την επιρροή στην οδική ασφάλεια των πλευρικών οχημάτων. Ωστόσο, τα πλευρικά οχήματα μπορούν να επηρεάσουν τον τρόπο οδήγησης και να συμβάλλουν στην πρόκληση σύγκρουσης, είτε άμεσα με την εμπλοκή των ίδιων στην σύγκρουση, όπου το όχημα συγκρούεται με το πλευρικό, είτε έμμεσα επηρεάζοντας τον οδηγό λόγω της τοποθέτησής τους σε σχέση με το εξεταζόμενο όχημα.

Ο αλγόριθμος είναι σε θέση να **βελτιώσει την οδηγική συμπεριφορά**, καθώς, με βάση τα πραγματικά δεδομένα των οδηγών, το μοντέλο θα απέφυγε κρίσιμες καταστάσεις όπως προκύπτει από την ανταμοιβή που πήρε ο πράκτορας κατά τον έλεγχο του τελικού αλγορίθμου. Έτσι αν τα οχήματα βρίσκονταν σε κρίσιμες καταστάσεις, ο αλγόριθμος θα τις αναγνώριζε και θα επέλεγε ενέργειες οι οποίες θα τις απέτρεπαν.

Η **ενισχυτική μάθηση αποτελεί κατάλληλη μέθοδο για την διαμόρφωση των αλγορίθμων για τα συστήματα υποβοήθησης οδηγού** και τα συστήματα αυτόνομης οδήγησης. Οι μέθοδοι αυτοί είναι σε θέση να διαχειριστούν περιβάλλοντα τα οποία δεν είναι σαφώς ορισμένα και ο αριθμός των καταστάσεων στις οποίες μπορούν να βρεθούν

είναι άπειρος καθώς αποτελούνται από πολλές διαστάσεις. Η οδήγηση αποτελεί μια πολύπλοκη διεργασία κατά την οποία πρέπει να ληφθούν υπόψιν διάφοροι παράγοντες για την επιλογή της βέλτιστης ενέργειας. Συνεπώς, οι αλγόριθμοι ενισχυτικής μάθησης είναι σε θέση να αντιμετωπίσουν τέτοιες διεργασίες όπως προκύπτει και από τα αποτελέσματα της παρούσας Διπλωματικής Εργασίας.

Ο αλγόριθμος κατά την διάρκεια του ελέγχου έκανε **4 λάθη στις 54976 δοκιμές**, τα οποία οφείλονται στο ότι ο αλγόριθμος κατά την εκπαίδευση δεν εκτέθηκε σε αυτές τις καταστάσεις επαρκώς έτσι ώστε να εντοπίσει την ιδανική ενέργεια. Τα λάθη είναι πολύ λίγα, καθώς ο αλγόριθμος έχει εκπαιδευτεί με χρήση μιας μεγάλης βάσης δεδομένων, συνεπώς έχει εκτεθεί σε ποικίλες καταστάσεις και έχει μάθει να τις διαχειρίζεται. Ένα σύστημα υποβοήθησης οδηγού θα πρέπει να παρουσιάζει ελάχιστα λάθη, καθώς τυχόν λάθος ενέργειες μπορεί να οδηγήσουν σε σύγκρουση. Η αξιοπιστία τέτοιων συστημάτων πρέπει να είναι υψηλή για να μπορέσουν να γίνουν δεκτά και να χρησιμοποιηθούν.

Ο τελικός Πίνακας-Q περιέχει μερικά μηδενικά. Αυτό μπορεί να οφείλεται στο γεγονός ότι κατά την διάρκεια της εκπαίδευσης ο αλγόριθμος δεν εκτέθηκε σε ικανοποιητικό βαθμό σε αυτές τις καταστάσεις έτσι ώστε να δοκιμάσει τις διάφορες ενέργειες μέχρις ότου βρει την βέλτιστη που θα επιφέρει την μέγιστη ανταμοιβή. Ορισμένες καταστάσεις έχουν μηδενικά σε όλες τις ενέργειες. Αυτές οι καταστάσεις είναι αρκετά ακραίες και ο αλγόριθμος δεν είχε την ευκαιρία να εκτεθεί σε πολλές κρίσιμες καταστάσεις. Κατά την διάρκεια συλλογής των δεδομένων δεν παρατηρήθηκαν πολλές κρίσιμες καταστάσεις ή οριακές συγκρούσεις για να μάθει ο αλγόριθμος από αυτές και να επεκτείνει το εύρος των καταστάσεων το οποίο μπορεί να διαχειριστεί. Συνεπώς ο Πίνακας-Q περιέχει μηδενικά σε αυτές τις καταστάσεις και ο αλγόριθμος δεν θα ήταν σε θέση να επιλέξει την σωστή ενέργεια σε περίπτωση που βρισκόταν σε αυτές τις καταστάσεις.

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ.....	1
1.1 ΓΕΝΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	1
1.2 ΣΚΟΠΟΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	3
1.3 ΔΙΑΡΘΡΩΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	4
2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ.....	6
2.1 ΕΙΣΑΓΩΓΗ	6
2.2 ΣΥΝΑΦΕΙΣ ΕΡΕΥΝΕΣ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΕΣ	6
2.3 ΣΥΝΟΨΗ.....	10
3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ.....	12
3.1 ΕΙΣΑΓΩΓΗ	12
3.2 ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ	12
3.3 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	14
3.3.1 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ	14
3.3.2 ΕΝΙΣΧΥΤΙΚΗ ΜΑΘΗΣΗ.....	14
3.3.3 ΟΙ ΕΞΙΣΩΣΕΙΣ BELLMAN	19
3.3.4 ΜΕΘΟΔΟΙ ΧΩΡΙΣ ΜΟΝΤΕΛΟ.....	20
3.3.5 ΕΞΕΡΕΥΝΗΣΗ ΕΝΑΝΤΙ ΕΚΜΕΤΑΛΛΕΥΣΗΣ.....	21
3.3.6 ΡΥΘΜΟΣ ΕΚΜΑΘΗΣΗΣ	22
4. ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	23
4.1 ΕΙΣΑΓΩΓΗ	23
4.2 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ.....	23
4.3 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	25
4.4 ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ	29
5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	32
5.1 ΕΙΣΑΓΩΓΗ	32
5.2 ΚΑΤΑΣΤΑΣΕΙΣ.....	32
5.3 ΕΝΕΡΓΕΙΕΣ.....	39
5.4 ΣΥΝΑΡΤΗΣΗ ΑΝΤΑΜΟΙΒΗΣ.....	39
5.5 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ	41
5.6 ΑΠΟΤΕΛΕΣΜΑΤΑ	41

6. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ	51
6.1 ΕΙΣΑΓΩΓΗ	51
6.2 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ	51
6.3 ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑ	53
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	55

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1: Διάγραμμα ροής διεργασιών Διπλωματικής Εργασίας	5
Διάγραμμα 2: Διαδικασία ανάπτυξης ενός συστήματος ενισχυτικής μάθησης από δεδομένα συλλεγμένα από ΣμηΕΑ	13
Διάγραμμα 3: Διασπορά των αποστάσεων του εξεταζόμενου με το αριστερά όχημα	29
Διάγραμμα 4: Διασπορά των αποστάσεων του εξεταζόμενου με το δεξιά όχημα	30
Διάγραμμα 5: Διασπορά των αποστάσεων του εξεταζόμενου με το μπροστά όχημα	31
Διάγραμμα 6: Διασπορά των αποστάσεων του εξεταζόμενου με το πίσω όχημα	31
Διάγραμμα 7: Έλεγχος μπροστά-πίσω οχημάτων	38
Διάγραμμα 8: Έλεγχος πλευρικής κυκλοφορίας	38
Διάγραμμα 9: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $l_r=0.2$	42
Διάγραμμα 10: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $l_r=0.5$	42
Διάγραμμα 11: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $\gamma=0.7$	43
Διάγραμμα 12: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $\gamma=0.9$	44
Διάγραμμα 13: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $e=0.5$	45
Διάγραμμα 14: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $e=0.7$	45

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1: Δομή των δεδομένων τροχιών	26
Πίνακας 2: Επιτρεπτή απόκλιση αποστάσεων	28
Πίνακας 3: Πίνακας-Q	49

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1: Αισθητήρες οχημάτων (Πηγή: Cooper 2017)	2
Εικόνα 2: Αρχιτεκτονική Συνελικτικού Νευρωνικού Δικτύου (Πηγή: Yu et al. 2016)	8
Εικόνα 3: Διαγραμματική απεικόνιση της κυκλοφοριακής εικόνας κατά την διάρκεια του cut-in (Πηγή: Rui Fu et al. 2019)	8
Εικόνα 4: Εννοιολογικό διάγραμμα του μοντέλου ακολουθούντος οχήματος με μεθόδους βαθιάς ενισχυτικής μάθησης (Πηγή: Zhu et al. 2018)	9
Εικόνα 5: Τύποι μηχανικής μάθησης (Πηγή: Shaikh 2017)	15
Εικόνα 6: Η Διαδικασία της ενισχυτικής μάθησης	16
Εικόνα 7: Αλυσίδα Markov (Πηγή: Silver 2015)	17
Εικόνα 8: Ο αλγόριθμος Q-Μάθησης (Πηγή: Choudhary 2019)	21
Εικόνα 9: Σημεία τοποθέτησης ΣμηΕΑ (Πηγή: Barmounakis et al. 2019)	24
Εικόνα 10: Αναπαράσταση των διαφόρων κινήσεων σε μια εκτεταμένη περιοχή (Πηγή: Barmounakis et al. 2019)	25
Εικόνα 11: Χρόνος για σύγκρουση μεταξύ δυο οχημάτων (Πηγή: Minderhoud et al. 2000)	33
Εικόνα 12: Κανονική Κατανομή (Πηγή: Galarnyk 2018)	36
Εικόνα 13: Έννοια της μετανόησης (Πηγή: Proutiere et al. 2019)	46

1. ΕΙΣΑΓΩΓΗ

1.1 ΓΕΝΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Στη σημερινή εποχή, οι **οδικές μεταφορές** αποτελούν ένα αναπόσπαστο κομμάτι της καθημερινότητας και συμβάλουν στην ευημερία, την ανάπτυξη και την πρόοδο της κοινωνίας. Η πλειοψηφία των μετακινήσεων γίνεται με χρήση ιδιωτικού οχήματος, των οποίων οι πωλήσεις τις τελευταίες δεκαετίες έχουν αυξηθεί σημαντικά, λόγω της μείωσης του κόστους των οχημάτων και της ταυτόχρονης αύξησης του κατά κεφαλήν εισοδήματος. Η μείωση του κόστους οφείλεται στην σημαντική πρόοδο της βιομηχανίας των αυτοκινήτων και στον έντονο ανταγωνισμό που υπάρχει μεταξύ των διάφορων εταιριών.

Η αύξηση της χρήσης του ΙΧ σχετίζεται εμμέσως με την αύξηση των οδικών ατυχημάτων. Τα οδικά ατυχήματα αποτελούν την 8^η κυριότερη αιτία θανάτου παγκοσμίως με **1.35 εκατομμύρια καταγεγραμμένους θανάτους** (World Health Organization, 2018), ενώ αποτελούν την πρώτη αιτία θανάτου σε ηλικίες 5-29 ετών. Ταυτόχρονα έχουν τεράστιες κοινωνικές και οικονομικές συνέπειες, καθώς το εκτιμώμενο κόστος των νεκρών, των τραυματιών και των υλικών ζημιών των καταγεγραμμένων οδικών ατυχημάτων στην Ελλάδα ξεπερνά τα 3 δις ευρώ. Είναι σκόπιμο λοιπόν, να καταβληθούν προσπάθειες για τη βελτίωση της οδικής ασφάλειας και την μείωση του αριθμού των ατυχημάτων.

Για το 95% των ατυχημάτων ευθύνεται εν μέρει ή εξολοκλήρου ο άνθρωπος (Φραντζεσκάκης et al. 1994). Επομένως, θα πρέπει να γίνουν προσπάθειες να **μειωθεί η συμμετοχή του ανθρώπινου παράγοντα** στην διαδικασία της οδήγησης, καθώς μία τέτοια μείωση θα οδηγήσει και σε αύξηση του επιπέδου της οδικής ασφάλειας. Όσον αφορά στην διαδικασία της οδήγησης, ο άνθρωπος δεν είναι αξιόπιστος. Αυτό συμβαίνει διότι, ο κάθε οδηγός έχει διαφορετική αντιληπτική ικανότητα οδήγησης, διαφορετικό χρόνο αντίδρασης, διαφορετική οξύτητα όρασης και λοιπά. Ταυτόχρονα πολλές φορές ο άνθρωπος δεν τηρεί τους κανόνες που διέπουν την κυκλοφορία της εκάστοτε περιοχής, αγνοώντας τον Κώδικα Οδικής Κυκλοφορίας και θέτοντας σε κίνδυνο τόσο την δική του ζωή όσο και αυτής των υπόλοιπων χρηστών της οδού. Για τη βελτίωση της οδικής ασφάλειας στο παρελθόν είχαν εφαρμοσθεί ορισμένα μέτρα παθητικής προστασίας, όπως ζώνες, αερόσακοι και γυαλιά ανθεκτικά στη θραύση. Αυτά τα μέτρα συμβάλουν στη μείωση των επιπτώσεων τυχόν συγκρούσεων. Σήμερα, με την εξέλιξη της τεχνολογίας γίνεται προσπάθεια να εισαχθούν στα οχήματα ορισμένα συστήματα ενεργητικής προστασίας τα οποία **αποτρέπουν εξολοκλήρου την σύγκρουση**.

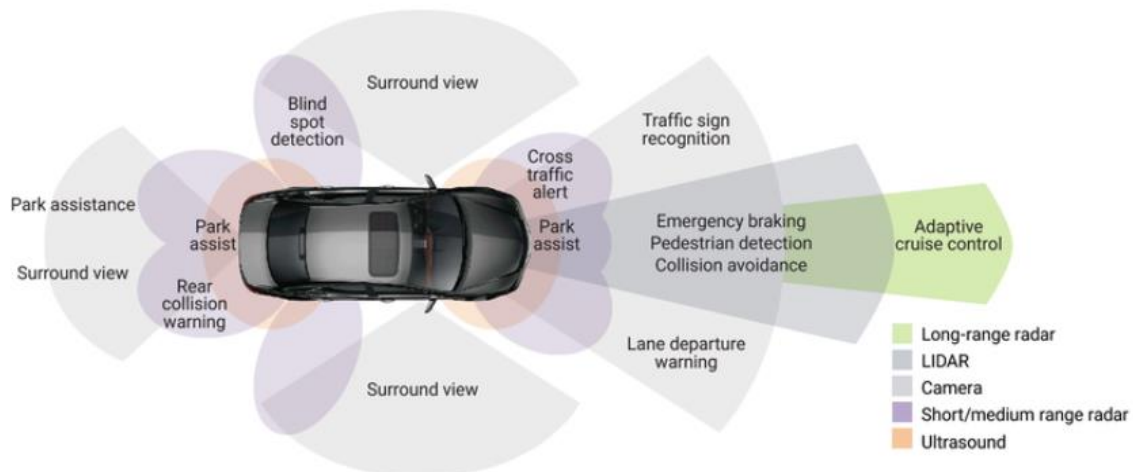
Συστήματα Υποβοήθησης Οδηγού

Η μείωση των οδικών ατυχημάτων μπορεί να επιτευχθεί με την **εισαγωγή συστημάτων υποβοήθησης οδήγησης** εντός του οχήματος. Η μετάβαση στην αυτόνομη κυκλοφορία κατά την οποία ο πλήρης στόλος των οχημάτων θα κινείται αυτόνομα (επίπεδο αυτοματισμού 5) γίνεται σταδιακά, καθώς προοδεύει η αντίστοιχη τεχνολογία συστημάτων

αυτοματισμού, υποβοήθησης και καθοδήγησης των οδηγών καθώς και οι αλγόριθμοι αυτόνομης οδήγησης. Τα συστήματα αυτά είναι αρκετά πιο αξιόπιστα από τον άνθρωπο και κάνουν πολύ λιγότερα λάθη, συμβάλλοντας έτσι σε ένα ασφαλέστερο περιβάλλον οδήγησης. Έτσι, εισάγονται όλο και περισσότερα συστήματα υποβοήθησης οδηγού (**ADAS**) στα νέα οχήματα που κυκλοφορούν στην αγορά. Τα συστήματα αυτά αποσκοπούν στη βελτίωση της ασφάλειας των οχημάτων και γενικότερα τη βελτίωση της οδικής ασφάλειας. Μερικά από αυτά τα συστήματα είναι:

- Σύστημα προσαρμογής ταχύτητας ταξιδιού (Adaptive cruise control)
- Σύστημα υποβοήθησης πέδησης (Brake Assist)
- Σύστημα αποφυγής σύγκρουσης (Collision Avoidance)
- Σύστημα ανίχνευσης υπνηλίας οδηγού (Drowsy driver detection system)
- Σύστημα αλλαγής λωρίδας (Lane changing assistant)
- Σύστημα διατήρησης λωρίδας (Lane keeping assistant)
- Σύστημα ενεργητικής προστασίας πεζών (Active pedestrian protection)
- Σύστημα αντιμπλοκαρίσματος τροχών (Anti-lock brakes)
- Σύστημα νυχτερινής όρασης (Night vision)
- Σύστημα αναγνώρισης οδικών σημάτων (Traffic sign recognition)
- Σύστημα ανίχνευσης χαμηλής τριβής (Low friction detection)

Όλα αυτά τα συστήματα αποσκοπούν στην απλοποίηση της διαδικασίας της οδήγησης και στην αύξηση του επιπέδου ασφάλειας. Για την υλοποίηση τους απαιτούνται: α. ένα πλήθος αισθητήρων πάνω στο όχημα (Εικόνα 1) και β. η διαμόρφωση αλγορίθμων για την επεξεργασία και μοντελοποίηση των δεδομένων που συλλέγονται.



Εικόνα 1: Αισθητήρες οχημάτων (Πηγή: Cooper 2017)

Μια επιπλέον μέθοδος μηχανικής μάθησης είναι η **ενισχυτική μάθηση**, όπου το σύστημα προσπαθεί να μάθει απευθείας από το περιβάλλον και στόχος είναι η μεγιστοποίηση ενός αριθμητικού σήματος ενίσχυσης (ανταμοιβή). Το σύστημα δεν καθοδηγείται από κάποιον εξωτερικό επιβλέποντα για το ποια ενέργεια θα ακολουθήσει αλλά πρέπει να ανακαλύψει μόνο του ποιες ενέργειες είναι αυτές που θα του αποφέρουν το μεγαλύτερο κέρδος. Τα μοντέλα ενισχυτικής μάθησης απαιτούν μεγάλη υπολογιστική ικανότητα και μεγάλο όγκο δεδομένων για εκπαίδευση, οπότε μόλις τα τελευταία χρόνια έχουν καταφέρει να εξελιχθούν σε μεγάλο βαθμό. Χάρη στον τρόπο με τον οποίο λειτουργεί, η μέθοδος έχει εφαρμοσθεί σε πολλές κλάδους όπως στη ρομποτική, στα οικονομικά και στα συστήματα αυτόνομης οδήγησης.

Ιδιαίτερα τα τελευταία χρόνια, οι μέθοδοι ενισχυτικής μάθησης και τα νευρωνικά δίκτυα έχουν συνδυαστεί με σκοπό την δημιουργία αξιόπιστων και αποτελεσματικών μοντέλων ικανά να εφαρμοστούν σε πολύπλοκα προβλήματα, όπως αυτό της οδήγησης (Meixin et al. 2019).

1.2 ΣΚΟΠΟΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Σκοπός της παρούσας Διπλωματικής Εργασίας είναι η **δημιουργία αλγορίθμων ενισχυτικής μάθησης με στόχο τη βελτίωση της οδηγικής συμπεριφοράς**. Συγκεκριμένα, εξετάζεται η δυνατότητα των αλγορίθμων ενισχυτικής μάθησης να ανταποκριθούν στις ανάγκες ενός συστήματος υποβοήθησης οδηγού το οποίο θα δίνει μηνύματα στο οδηγό για την κίνηση του σε σχέση με την υπόλοιπη κυκλοφορία. Ο αλγόριθμος αυτός θα πρέπει να είναι σε θέση σε κάθε χρονική στιγμή να αναγνωρίζει το επίπεδο της οδικής ασφάλειας της κατάστασης στην οποία βρίσκεται το όχημα και να προτείνει ενέργειες που θα οδηγήσουν το όχημα σε μια ασφαλέστερη κατάσταση.

Ο αλγόριθμος προορίζεται για χρήση σε αυτοκίνητα, όπου θα λαμβάνει από τους κατάλληλους αισθητήρες πληροφορίες σχετικά με τις αποστάσεις του αυτοκινήτου από τα πλευρικά, το μπροστά και το πίσω όχημα καθώς και τις ταχύτητες του μπροστά και του πίσω οχήματος, θα αναγνωρίζει την κατάσταση στην οποία βρίσκεται το όχημα και θα προτείνει στον οδηγό κατάλληλες ενέργειες λαμβάνοντας υπόψιν τον χρόνο αντίδρασης του οδηγού από την στιγμή που θα λάβει το μήνυμα μέχρι να ενεργήσει.

Ο αλγόριθμος ενισχυτικής μάθησης που χρησιμοποιήθηκε είναι η **Q-Μάθηση**, ο οποίος είναι αλγόριθμος χωρίς τη χρήση μοντέλου και από τον οποίο προκύπτει ο Πίνακας-Q. Ο πίνακας αυτό περιέχει τις αξίες για όλους του συνδυασμούς καταστάσεων-ενεργειών και με βάση αυτόν ο αλγόριθμος επιλέγει τις κατάλληλες σε κάθε χρονική στιγμή ενέργειες. Συγκεκριμένα όσο μεγαλύτερη είναι η αξία της κάθε ενέργειας για μια δεδομένη κατάσταση τόσο πιο ελκυστική είναι αυτή η ενέργεια. Η ενέργεια με την μέγιστη αξία είναι αυτή που θα οδηγήσει στην ασφαλέστερη κατάσταση.

Τα **δεδομένα** για την εκπαίδευση του αλγορίθμου λήφθηκαν από πραγματικές τροχιές οχημάτων σε κεντρικούς άξονες της Αθήνας. Οι οδοί αυτοί αποτελούνται από τουλάχιστον τρεις λωρίδες κυκλοφορίας, έτσι ώστε να είναι δυνατή η ταυτόχρονη ύπαρξη οχημάτων

δεξιά και αριστερά του εξεταζόμενου. Με αυτό τον τρόπο ο αλγόριθμος εκτίθεται σε μια πληθώρα καταστάσεων που εμφανίζονται σε κεντρικά τμήματα μεγάλων πόλεων με υψηλούς κυκλοφοριακούς φόρτους. Οι τροχιές των οχημάτων εξήχθησαν από βίντεο με κάμερα τοποθετημένη σε ΣμηΕΑ τα οποία πετούσαν σε σταθερό σημείο πάνω από την εξεταζόμενη περιοχή.

1.3 ΔΙΑΡΘΡΩΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Η διαδικασία υλοποίησης της Διπλωματικής Εργασίας ακολουθεί την ροή του Διαγράμματος 1. Το πρώτο κεφάλαιο αποτελεί την **εισαγωγή** της Διπλωματικής Εργασίας και παρουσιάζει στον αναγνώστη το γενικότερο πλαίσιο της οδικής ασφάλειας και τις τεχνολογικές προόδους στον τομέα βελτίωσης της ασφάλειας των οχημάτων. Ξεκινά με μία αναφορά στο γενικότερο πρόβλημα της οδικής ασφάλειας στην Ευρώπη και στην Ελλάδα. Ύστερα, παρουσιάζονται τα διάφορα συστήματα υποβοήθησης οδηγού και οι κύριοι αλγόριθμοι που χρησιμοποιούνται για την λειτουργία αυτών των συστημάτων και τέλος, περιγράφεται ο στόχος της διπλωματικής εργασίας.

Στο δεύτερο κεφάλαιο παρουσιάζονται τα αποτελέσματα της **βιβλιογραφικής ανασκόπησης** όπως αυτά προέκυψαν από την αναζήτηση ερευνών με παρεμφερές αντικείμενο. Η αναζήτηση βιβλιογραφικών αναφορών αφορά στην χρήση κατάλληλων πηγών για τον εντοπισμό τίτλων εργασιών, οι οποίες έχουν συναφές θέμα με αυτό της Διπλωματικής Εργασίας, δηλαδή έρευνες πάνω στα συστήματα υποβοήθησης οδηγού και στους αλγορίθμους ενισχυτικής μάθησης. Παρατίθενται εργασίες που έχουν δημοσιευθεί σε συνέδρια, επιστημονικά περιοδικά, άρθρα η συγγράμματα. Κατόπιν, πραγματοποιείται σύνοψη των αποτελεσμάτων των ερευνών και κριτική αξιολόγηση τους, ώστε να διαπιστωθεί η ποιότητα των αποτελεσμάτων και η ικανότητά τους να συμβάλλουν στην παρούσα Διπλωματική Εργασία.

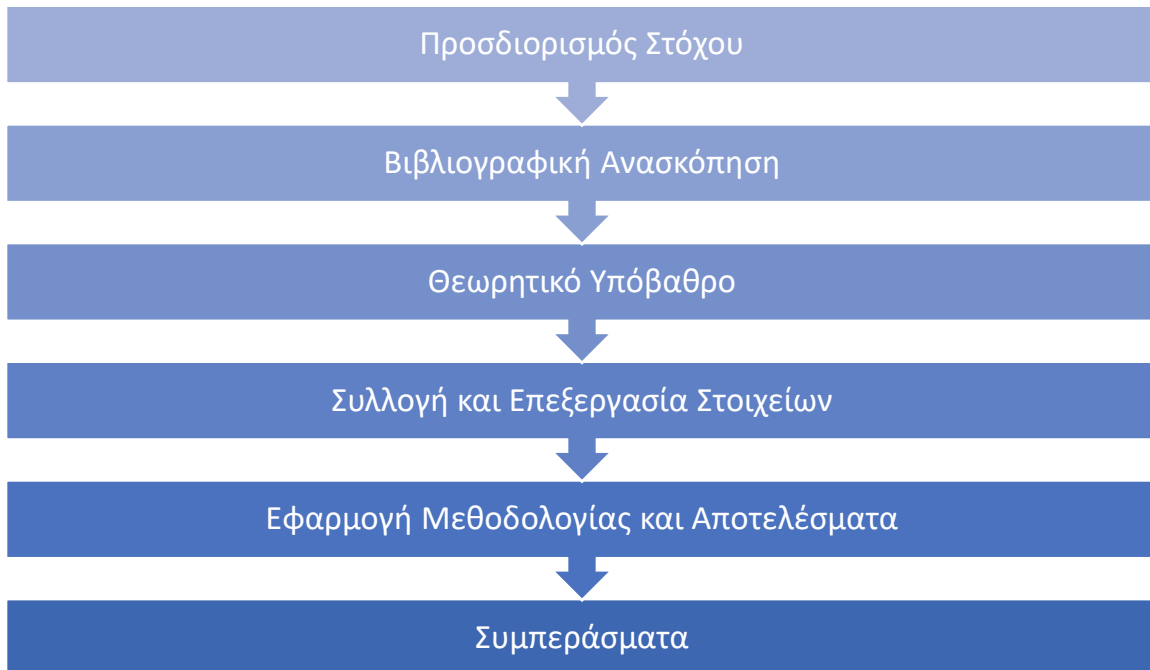
Στο τρίτο κεφάλαιο, το οποίο αποτελεί το **θεωρητικό υπόβαθρο** της Διπλωματικής Εργασίας, παρουσιάζεται και αναλύεται η επιλεγείσα μεθοδολογία. Δίνεται ιδιαίτερη έμφαση στις μαθηματικές σχέσεις και τις θεωρίες πάνω στις οποίες βασίζεται η ενισχυτική μάθηση, ενώ ταυτόχρονα περιγράφονται οι αρχές και οι τρόποι με τους οποίους εφαρμόζεται η μεθοδολογία.

Στο τέταρτο κεφάλαιο περιγράφεται η διαδικασία **συλλογής και επεξεργασίας των στοιχείων** που χρησιμοποιήθηκαν μετέπειτα στην εκπαίδευση του αλγορίθμου ενισχυτικής μάθησης. Αρχικά παρουσιάζονται οι πηγές των στοιχείων, ο τρόπος συλλογής και προκαταρκτικής επεξεργασίας και τα βασικά χαρακτηριστικά τους. Έπειτα αναλύεται ο τρόπος επεξεργασίας των στοιχείων για την διαμόρφωση της τελικής μορφής και την εισαγωγή τους στον αλγόριθμο.

Στο πέμπτο κεφάλαιο γίνεται η παρουσίαση της **διαδικασίας ανάπτυξης και εφαρμογής των τελικών αλγορίθμων**. Γίνεται αναφορά στον τρόπο με τον οποίο εφαρμόσθηκε η μεθοδολογία, αναλύονται οι παράμετροι της ενισχυτικής μάθησης και οι διάφοροι

συνδυασμοί τιμών. Στη συνέχεια παρουσιάζεται ο τελικός αλγόριθμος και ο Πίνακας-Q που προέκυψε ως μεταβλητή εξόδου του αλγόριθμου.

Στο έκτο κεφάλαιο πραγματοποιείται σύνοψη των αποτελεσμάτων, παρατίθενται τα **συμπεράσματα** που προέκυψαν από την ερμηνεία των αποτελεσμάτων και γίνεται αναφορά στον τρόπο με τον οποίο μπορούν να χρησιμοποιηθούν τα αποτελέσματα. Τέλος παρουσιάζονται προτάσεις για περαιτέρω έρευνα και εξέλιξη των συγκεκριμένων συστημάτων.



Διάγραμμα 1: Διάγραμμα ροής διεργασιών Διπλωματικής Εργασίας

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 ΕΙΣΑΓΩΓΗ

Στο παρόν κεφάλαιο πραγματοποιείται η βιβλιογραφική ανασκόπηση για τις ανάγκες της παρούσας Διπλωματικής Εργασίας. Για τον σκοπό αυτό, αναζητήθηκαν επιστημονικές έρευνες που αφορούν στα **συστήματα υποβοήθησης οδηγού και στις μεθόδους ενισχυτικής μάθησης**. Τα στοιχεία που εξετάζονται προκύπτουν από έρευνες που έχουν πραγματοποιηθεί στην Ελλάδα, αλλά και στο εξωτερικό. Για κάθε εργασία γίνεται συνοπτική αναφορά στη μεθοδολογία που ακολουθήθηκε και στα συμπεράσματα που προέκυψαν. Τέλος, με βάση τη βιβλιογραφική ανασκόπηση, προσδιορίστηκε το ακριβές αντικείμενο της Διπλωματικής Εργασίας και έγινε η καταλληλότερη επιλογή μεθοδολογίας.

2.2 ΣΥΝΑΦΕΙΣ ΕΡΕΥΝΕΣ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΕΣ

Η βιβλιογραφική έρευνα επιτρέπει τον εντοπισμό αντικειμένων τα οποία δεν έχουν καλυφθεί επαρκώς από ήδη υπάρχουσες έρευνες και τον έλεγχο των αποτελεσμάτων της παρούσας Διπλωματικής Εργασίας με αυτά της διεθνούς βιβλιογραφίας.

Τα μοντέλα ακολουθούντος οχήματος (car following) έχουν διερευνηθεί με διάφορες προσεγγίσεις. Παλαιότερα τα μοντέλα αυτά βασίζονταν κυρίως σε μαθηματικές σχέσεις και προέκυπταν από θεωρίες της κυκλοφοριακής ροής. Ωστόσο, τα τελευταία χρόνια έχει καταστεί δυνατή η συλλογή μεγάλων όγκων **κυκλοφοριακών δεδομένων τα οποία μπορούν να χρησιμοποιηθούν στα μοντέλα** ακολουθούντος οχήματος (Parathanasoroulou et al. 2015). Με αυτό τον τρόπο τα μοντέλα μπορούν να είναι πιο ευέλικτα και γίνεται εφικτή η ενσωμάτωση επιπλέον πληροφοριών στο μοντέλο. Το μοντέλο που προτείνεται στην έρευνα μπορεί να ενσωματωθεί και σε μοντέλα μικροσκοπικής προσομοίωσης κυκλοφορίας. Τα δεδομένα για το πείραμα λήφθηκαν από μια ακολουθία οχημάτων στην Νάπολη της Ιταλίας. Η έρευνα κατέληξε σε ένα μοντέλο το οποίο παρέχει πιο αξιόπιστα αποτελέσματα από αυτό του Girr που βασίζεται σε συμβατικές μεθόδους.

Επίσης ένα από τα συμπεράσματα της παραπάνω έρευνας είναι ότι οι **μέθοδοι μηχανικής μάθησης** παρέχουν μεγάλη ευελιξία και ταχύτητα στην διαχείριση των δεδομένων και **δεν απαιτούν έντονη βαθμονόμηση** των παραμέτρων, όπως τα παραδοσιακά μοντέλα ακολουθούντος οχήματος. Ωστόσο, τα παραδοσιακά μοντέλα έχουν το πλεονέκτημα ότι βασίζονται σε θεωρίες της κυκλοφοριακής ροής, πράγμα που τις κάνει κατανοητές από τον άνθρωπο, σε αντίθεση με τις υπολογιστικές προσεγγίσεις που δεν συνεισφέρουν στην εξέλιξη της κυκλοφοριακής θεωρίας.

Οι Peng Liao et al. (2019) προτείνουν ένα μοντέλο ακολουθούντος οχήματος για την εξασφάλιση της οδικής ασφάλειας, της σταθερότητας, της άνεσης και της εξοικονόμησης καυσίμων λαμβάνοντας υπόψιν τις διαφορετικές οδηγικές συνήθειες του κάθε οδηγού. Κατά την διάρκεια του φαινομένου του ακολουθούντος οχήματος, η επιτάχυνση του

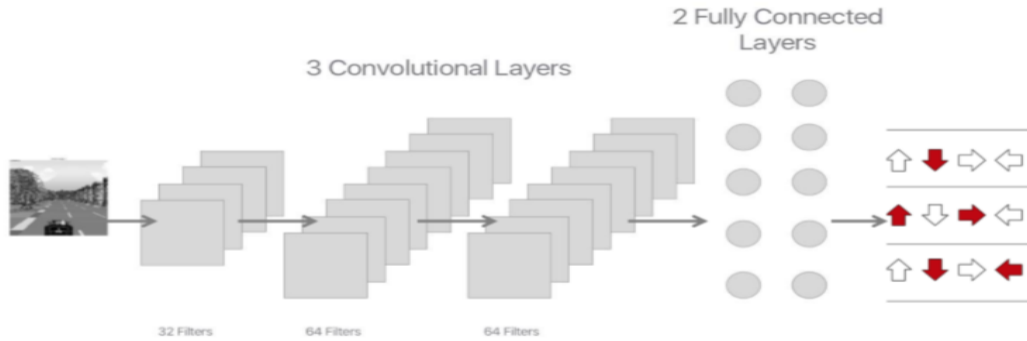
οχήματος που ακολουθεί καθορίζεται από την στρατηγική οδήγησης για διάφορες αποστάσεις από το μπροστινό όχημα και με βάση τις οδηγικές συνήθειες. Όταν η απόσταση είναι υπερβολικά μεγάλη, το όχημα που ακολουθεί μειώνει την διαφορά με μία βέλτιστη, με βάση την άνεση και την οικονομία καυσίμων, επιτάχυνση για να εξασφαλιστεί η σταθερή συμπεριφορά. Ομοίως όταν η απόσταση είναι πολύ μικρή, το όχημα που ακολουθεί φρενάρει για να διατηρηθεί η ασφάλεια. Όταν η απόσταση ανάμεσα στα δύο οχήματα είναι εντός του επιθυμητού εύρους, το όχημα που ακολουθεί μπορεί να διατηρήσει μια σχετικά σταθερή ταχύτητα. Τα αποτελέσματα δείχνουν ότι το προτεινόμενο μοντέλο μπορεί να διασφαλίσει της ασφάλεια, την σταθερότητα, την άνεση και την βέλτιστη κατανάλωση καυσίμων.

Το μοντέλο αυτό θεωρεί τρία εύρη αποστάσεων για την **ανάπτυξη τριών στρατηγικών οδήγησης**:

- Μια επικίνδυνη απόσταση για την αποφυγή συγκρούσεων και τη βελτίωση της ασφάλειας σε περίπτωση έκτακτου φρεναρίσματος. Αυτή η απόσταση διαφέρει για κάθε οδηγό και εξαρτάται από τα χαρακτηριστικά της συμπεριφοράς τους.
- Μία μεγάλη απόσταση πέραν της οποίας ο οδηγός του οχήματος που ακολουθεί πρέπει να επιταχύνει ώστε να μειώσει την απόσταση και στην συνέχεια να επιβραδύνει. Αυτή η απόσταση είναι ίση με 150 μέτρα για όλους τους οδηγούς.
- Μια επιθυμητή απόσταση η οποία είναι μεγαλύτερη από την επικίνδυνη απόσταση και μικρότερη από 150 μέτρα.

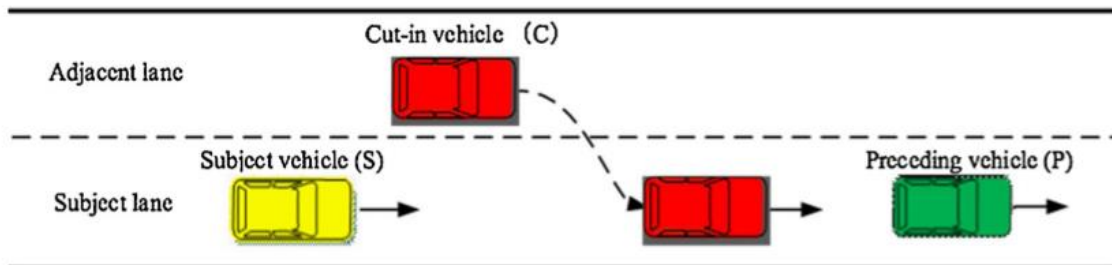
Με βάση αυτές τις αποστάσεις, την βέλτιστη από πλευράς κατανάλωσης καυσίμων επιτάχυνση και επιβράδυνση και την άνεση του κάθε οδηγού προκύπτει το μοντέλο. Ωστόσο, ένας περιορισμός του μοντέλου αυτού είναι η έλλειψη βαθμονόμησης με βάση πειραματικά/εμπειρικά δεδομένα και η αγνόηση τυχόν επικοινωνίας μεταξύ των οχημάτων.

Οι Yu et al. (2016) προσπάθησαν να ελέγξουν ένα **προσομοιωμένο αυτοκίνητο με χρήση βαθιάς ενισχυτικής μάθησης**. Αρχικά, έγινε χρήση Q-Δικτύου όπως αυτό προέκυψε από τους Mnih et al. (2015) όπου επιχειρήθηκε η δημιουργία ενός μοντέλου ικανό να συμπεριφερθεί όπως οι άνθρωποι. Οι πράκτορες του μοντέλου είναι σε θέση να αντλούν αποτελεσματικά παραστάσεις του περιβάλλοντος από εισροές πολλών διαστάσεων και να τις χρησιμοποιούν για τη γενίκευση της εμπειρίας του παρελθόντος σε νέες καταστάσεις. Στην συνέχεια, έγιναν προσπάθειες τροποποίησης του αλγορίθμου των Mnih et al. (2015) για να βελτιωθεί η απόδοση στην συγκεκριμένη διεργασία, μέσω της εισαγωγής διαφόρων συναρτήσεων ανταμοιβής, έτσι ώστε να επιτευχθεί συγκεκριμένη οδηγική συμπεριφορά, διπλής Q-Μάθησης και άλλων υπερπαραμέτρων (Εικόνα 2). Στο τελικό μοντέλο ο πράκτορας έμαθε ικανοποιητικά να στρίβει για να αποφύγει τις συγκρούσεις, ωστόσο δεν κατάφερε να μάθει να αποφεύγει εμπόδια, όπως άλλα οχήματα, το οποίο μπορεί να οφείλεται στην ανάγκη για περαιτέρω εκπαίδευση.



Εικόνα 2: Αρχιτεκτονική Συνελκτικού Νευρωνικού Δικτύου (Πηγή: Yu et al. 2016)

Τα περισσότερα μοντέλα ακολουθούντος οχήματος για αυτόνομα οχήματα και για συνδεδεμένα οχήματα δεν λαμβάνουν υπόψιν την είσοδο στην μέση άλλων οχημάτων οπότε και **κόβεται η ροή του οχήματος που ακολουθεί** (Εικόνα 3). Οι Rui Fu et al. (2019) προτείνουν ένα μοντέλο για αυτόνομα οχήματα το οποίο εξετάζει τις επιπτώσεις αυτού του φαινομένου και τις προσδοκίες των οδηγών. Διερευνήθηκαν η θέση εισχώρησης του τρίτου οχήματος, ο χρόνος αντίδρασης, η επιτάχυνση και η ιδανική απόσταση σε έναν πραγματικό αυτοκινητόδρομο με την χρήση ενός οχήματος με τα κατάλληλα όργανα. Τα οχήματα που εισέρχονται ενδιάμεσα στα οχήματα του φαινομένου του ακολουθούντος οχήματος διατηρούν μια ασφαλή απόσταση από το όχημα που προπορεύεται και μια μεγαλύτερη απόσταση από το όχημα που ακολουθεί για την αποφυγή σύγκρουσης. Για την πραγματοποίηση της έρευνας χρησιμοποιήθηκαν 46 μη-επαγγελματίες οδηγοί και 519 γεγονότα car following και 219 cut-in καταγράφηκαν.

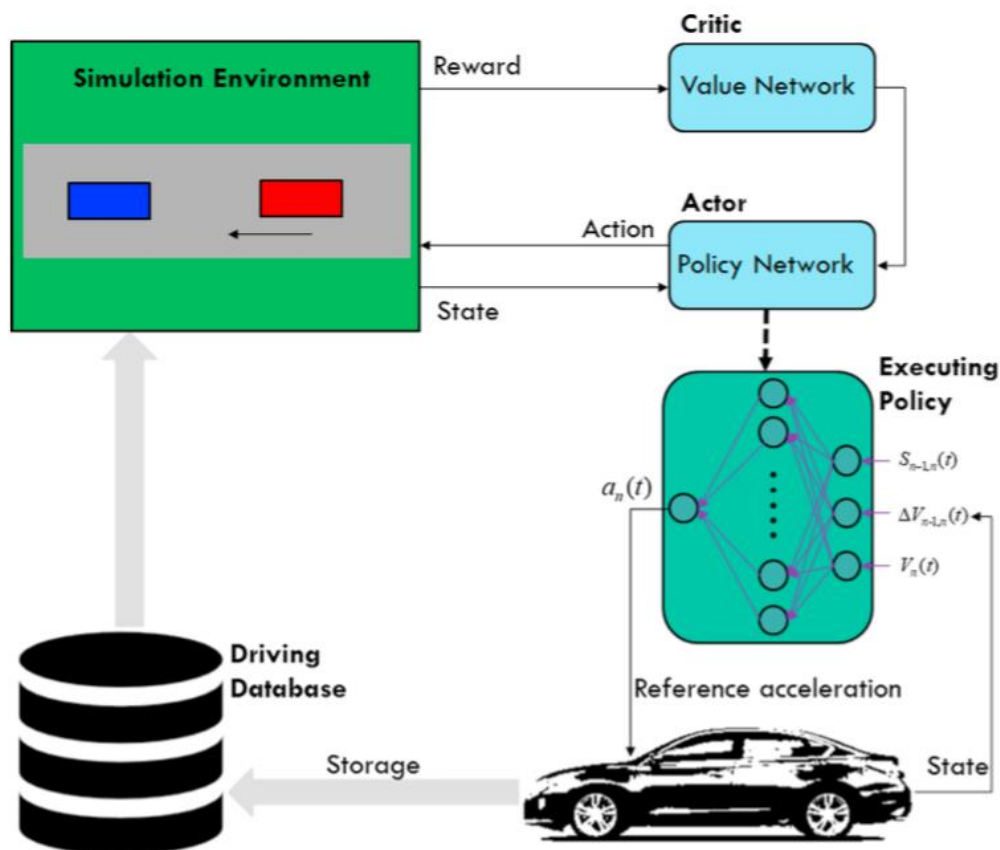


Εικόνα 3: Διαγραμματική απεικόνιση της κυκλοφοριακής εικόνας κατά την διάρκεια του cut-in (Πηγή: Rui Fu et al. 2019)

Η ανάλυση της συμπεριφοράς του οδηγού που ακολουθεί το μπροστινό όχημα, δείχνει ότι στα στάδια διατήρησης, ο χρόνος αντίδρασης αφότου εισέλθει το ενδιάμεσο όχημα είναι 0.85 δευτερόλεπτα για το ερέθισμα επιτάχυνσης και 0.7 δευτερόλεπτα για το ερέθισμα επιβράδυνσης. Αυτοί οι χρόνοι είναι μικρότεροι από τον χρόνο απόκρισης πριν την είσοδο του τρίτου οχήματος για την επιτάχυνση (1.95 δευτερόλεπτα) και για την επιβράδυνση (1.66 δευτερόλεπτα). Η επιτάχυνση και η επιθυμητή απόσταση είναι μικρότερες σε σχέση

με πριν την διακοπή του φαινομένου του ακολουθούντος οχήματος. Το μοντέλο που προτείνεται προσομοιάζει την ανθρώπινη συμπεριφορά και έχει μικρότερο χρόνο απόκρισης και μικρότερη επιβράδυνση σε τέτοιες καταστάσεις, ενώ μπορεί να χρησιμοποιηθεί για τη βελτίωση της ασφάλειας του ακολουθούντος οχήματος, της άνεσης του οδηγού και της εμπιστοσύνης στα αυτόνομα και συνδεδεμένα οχήματα.

Οι Zhua et al. (2019) προτείνουν ένα πλαίσιο για **αυτόνομη οδήγηση όπου το όχημα θα συμπεριφέρεται σαν άνθρωπος** (Εικόνα 4). Το μοντέλο επικεντρώνεται σε ακολουθούντα οχήματα και αναπτύσσεται με διάφορες μεθόδους ενισχυτικής μάθησης. Ιστορικά δεδομένα εισάγονται σε ένα περιβάλλον προσομοίωσης όπου ένας πράκτορας ενισχυτικής μάθησης μαθαίνει με δοκιμές και λαμβάνοντας μια ανταμοιβή με βάση το πόσο απέχει από τα εμπειρικά δεδομένα. Μια βέλτιστη πολιτική προκύπτει από αυτή την εκπαίδευση όπου ο πράκτορας είναι σε θέση να επιλέξει την ταχύτητα, την σχετική ταχύτητα σε σχέση με το όχημα που προπορεύεται, την ενδιάμεση απόσταση και την επιτάχυνση κατά την διάρκεια του φαινομένου του ακολουθούντος οχήματος. Τα δεδομένα του μοντέλου λήφθηκαν από την έρευνα “2015 Shanghai Naturalistic Driving Study” και περιείχαν δύο χιλιάδες περιόδους.



Εικόνα 4: Εννοιολογικό διάγραμμα του μοντέλου ακολουθούντος οχήματος με μεθόδους βαθιάς ενισχυτικής μάθησης (Πηγή: Zhu et al. 2018)

Τα αποτελέσματα της έρευνας δείχνουν ότι ένα μοντέλο βαθιάς ντετερμινιστικής βαθμιδωτής πολιτικής μπορεί να αναπαράγει την ανθρώπινη συμπεριφορά με μεγαλύτερη ακρίβεια σε σύγκριση με τα βασικά πρόσφατα μοντέλα, συμπεριλαμβανομένων του μοντέλου έξυπνου οδηγού, των μοντέλων που βασίζονται σε τοπικά σταθμισμένη παλινδρόμηση και τα συμβατικά νευρωνικά δίκτυα. Συγκεκριμένα, το μοντέλο έχει σφάλμα αποστάσεων 18% και σφάλμα ταχυτήτων 5%, μικρότερες τιμές από τις αντίστοιχες των προαναφερθέντων μοντέλων. Η έρευνα καταλήγει στο συμπέρασμα ότι η ενισχυτική μάθηση μπορεί να προσφέρει πληροφορίες για την συμπεριφορά οδηγών και μπορεί να συνεισφέρει στην ανάπτυξη αλγορίθμων αυτόνομης οδήγησης με συμπεριφορά παρόμοια με των ανθρώπων και μοντέλων κυκλοφοριακής ροής, ενώ έχει καλή ικανότητα γενίκευσης. Το μοντέλο αυτό μπορεί να χρησιμοποιηθεί στην μετάβαση στην πλήρως αυτόνομη κυκλοφορία όπου τα αυτόνομα οχήματα θα πρέπει να συμπεριφέρονται ως άνθρωποι για να υπάρχει ομοιομορφία και να μπορούν οι άνθρωποι να τα εμπιστευτούν και οι γύρω οδηγοί να αναμένουν τις ενέργειες, ώστε να μπορούν να αλληλοεπιδράσουν με αυτά με έναν φυσικό τρόπο.

2.3 ΣΥΝΟΨΗ

Στο κεφάλαιο αυτό παρουσιάστηκαν τα ευρήματα της βιβλιογραφικής ανασκόπησης που έγινε για τις ανάγκες της παρούσας Διπλωματικής Εργασίας. Πραγματοποιήθηκε ανάλυση ερευνών σχετικών με το φαινόμενο του ακολουθούντος οχήματος και των μεθοδολογιών που χρησιμοποιούνται για την ανάπτυξη των μοντέλων. Μερικά από τα βασικότερα συμπεράσματα που προκύπτουν από τις παραπάνω έρευνες είναι:

- Τα μοντέλα που βασίζονται σε πραγματικά δεδομένα μπορούν να οδηγήσουν σε καλύτερα αποτελέσματα σε σχέση με τα παραδοσιακά μοντέλα που βασίζονται σε μαθηματικές σχέσεις της κυκλοφοριακής ροής. Τα τελευταία χρόνια έχει καταστεί δυνατή η συλλογή και η επεξεργασία μεγάλων όγκων δεδομένων τα οποία μπορούν να χρησιμοποιηθούν για την δημιουργία μοντέλων με πολλές δυνατότητες.
- Οι μέθοδοι μηχανικής μάθησης παρέχουν την δυνατότητα ανάπτυξης πολύπλοκων μοντέλων με χρήση μεγάλων βάσεων δεδομένων. Η υπολογιστική δύναμη των σημερινών συστημάτων και η ευελιξία των αλγορίθμων μηχανικής μάθησης παρέχουν την δυνατότητα διαχείρισης μεγάλων όγκων δεδομένων σε πραγματικό χρόνο για την δημιουργία μοντέλων για διάφορα πολύπλοκα προβλήματα.
- Τα μοντέλα που λαμβάνουν υπόψιν την συμπεριφορά του οδηγού είναι πιθανό να εμφανίσουν μεγαλύτερη ασφάλεια, σταθερότητα, άνεση και βέλτιστη κατανάλωση καυσίμων.
- Τα μοντέλα ακολουθούντος οχήματος που λαμβάνουν υπόψιν την είσοδο ανάμεσα στα εξεταζόμενα οχήματα ενός τρίτου οχήματος οδηγούν σε αύξηση της οδικής ασφάλειας. Οι χρόνοι αντίδρασης του οδηγού που ακολουθεί κατά την διάρκεια εισόδου του τρίτου οχήματος είναι μικρότερη σε σύγκριση με αυτούς κατά την διάρκεια του φαινομένου του ακολουθούντος οχήματος.

- Οι μέθοδοι ενισχυτικής μάθησης μπορούν να χρησιμοποιηθούν αποτελεσματικά για την διαμόρφωση συστημάτων αυτόνομης οδήγησης. Τα προβλήματα που παρουσιάζουν πολλές διαστάσεις και περιβάλλον που δύσκολα προσδιορίζεται με μετρήσιμο αριθμό καταστάσεων είναι αρκετά δύσκολο να προσεγγισθούν με τυπικές μεθόδους. Η ενισχυτική μάθηση παρέχει τα εργαλεία για την επίλυση τέτοιου είδους προβλημάτων.
- Κατά την μετάβαση στην αυτόνομη κυκλοφορία τα οχήματα θα πρέπει να συμπεριφέρονται ως άνθρωποι έτσι ώστε να είναι αναμενόμενες οι ενέργειες τους από τους γύρω οδηγούς και να τα εμπιστεύονται με ευκολία οι χρήστες τους.

Συνοψίζοντας, η ενισχυτική μάθηση αποτελεί μια μέθοδο που τα τελευταία χρόνια έχει εξελιχθεί σε μεγάλο βαθμό για την ανάπτυξη μοντέλων για συστήματα οδήγησης. Αυτά τα μοντέλα ιδανικά θα πρέπει να εκπαιδεύονται με πραγματικά δεδομένα από την κυκλοφορία των αυτοκινήτων. Επιπλέον, δεν έχουν διερευνηθεί σε ικανοποιητικό βαθμό οι πλευρικές αποστάσεις και το πώς αυτές επηρεάζουν την ασφάλεια της οδήγησης, ιδιαίτερα σε συνδυασμό με τα μοντέλα ακολουθούστος οχήματος.

3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

3.1 ΕΙΣΑΓΩΓΗ

Στο παρόν κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο πάνω στο οποίο βασίστηκε η ανάλυση των στοιχείων και η εφαρμογή της μεθόδου της παρούσας Διπλωματικής Εργασίας. Αρχικά, παρουσιάζεται η μεθοδολογική προσέγγιση. Στη συνέχεια, γίνεται μια περιγραφή του προβλήματος της ενισχυτικής μάθησης και των διεργασιών Markov, ενώ παρουσιάζονται η έννοια της πολιτικής, η συνάρτηση ανταμοιβής, η συνάρτηση αξίας και οι εξισώσεις Bellman. Τέλος γίνεται αναφορά στις μεθόδους χωρίς μοντέλα, στην μέθοδο Q-Μάθησης που χρησιμοποιήθηκε για την ανάλυση των δεδομένων και στους όρους εξερεύνηση, εκμετάλλευση και ρυθμός εκμάθησης.

3.2 ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ

Η προτεινόμενη προσέγγιση για την ανάπτυξη ενός συστήματος ενισχυτικής μάθησης από δεδομένα συλλεγμένα από ΣμηΕΑ φαίνεται στο Διάγραμμα 2.

Η προσέγγιση περιλαμβάνει 3 βήματα:

Βήμα 1: Προετοιμασία δεδομένων

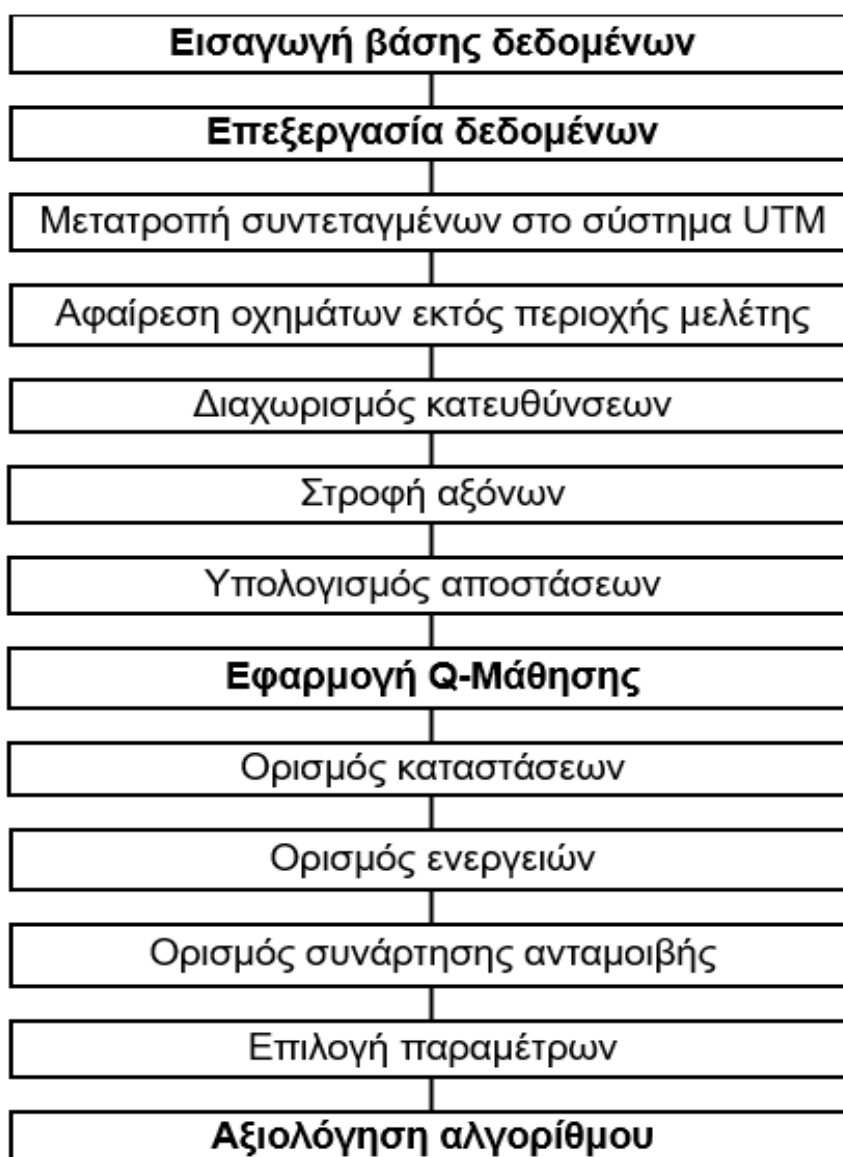
Τα δεδομένα περιλαμβάνουν τις συντεταγμένες των οχημάτων στο σύστημα WGS84 καθώς και τις ταχύτητες και επιταχύνσεις τους. Για την εισαγωγή τους στον αλγόριθμο Q-Μάθησης θα πρέπει να τροποποιηθούν κατάλληλα. Αρχικά γίνεται μετατροπή των συντεταγμένων στο προβολικό σύστημα UTM και αφαιρούνται όσα οχήματα βρίσκονται εκτός της περιοχής ελέγχου. Στην συνέχεια, με βάση το διάνυσμα κατεύθυνσης του κάθε οχήματος γίνεται διαχωρισμός των δύο κατευθύνσεων της κυκλοφορίας και οι συντεταγμένες τροποποιούνται, έτσι ώστε να ανταποκρίνονται στους άξονες της οδού. Τέλος, υπολογίζονται οι μικρότερες πλευρικές αποστάσεις μεταξύ του εξεταζόμενου και των γύρω οχημάτων.

Βήμα 2: Ανάπτυξη αλγορίθμου Q-learning

Για την ανάπτυξη του αλγορίθμου Q-Μάθησης, είναι απαραίτητος ο καθορισμός των καταστάσεων του περιβάλλοντος. Με βάση τον χρόνο για σύγκρουση του εξεταζόμενου με το μπροστά όχημα, του πίσω με το εξεταζόμενο όχημα καθώς και με βάση τις πλευρικές αποστάσεις καθορίζονται όλες οι πιθανές καταστάσεις. Οι ενέργειες που μπορεί να επιλέξει ο πράκτορας είναι όλοι οι συνδυασμοί ενεργειών που αφορούν τον έλεγχο του οχήματος, ενώ η συνάρτηση ανταμοιβής καθορίζει ποιες ενέργειες επιφέρουν θετικά αποτελέσματα.

Βήμα 3: Αξιολόγηση αλγορίθμου

Ο τελικός Πίνακας-Q που προκύπτει από τον αλγόριθμο αξιολογείται με βάση το διάγραμμα της αξίας της μέσης μέγιστης ενέργειας όλων των καταστάσεων το οποίο πρέπει να συγκλίνει σε μια τιμή. Η σύγκλιση αυτή αποτελεί ένδειξη ότι η εκπαίδευση του αλγορίθμου έχει φτάσει σε ικανοποιητικό επίπεδο. Ένας άλλος τρόπος αξιολόγησης του αλγορίθμου είναι ο έλεγχος με χρήση δεδομένων στα οποία δεν έχει εκτεθεί ο αλγόριθμος. Έτσι αναλόγως με τον αριθμό των λαθών που θα κάνει ο αλγόριθμος, εξάγονται συμπεράσματα σχετικά με την ποιότητά του. Τέλος, ελέγχεται η μετανόηση του αλγορίθμου που ισούται με την διαφορά της αθροιστικής ανταμοιβής της βέλτιστης πολιτικής από την αθροιστική ανταμοιβή της πολιτικής που ακολουθήθηκε.



Διάγραμμα 2: Διαδικασία ανάπτυξης ενός συστήματος ενισχυτικής μάθησης από δεδομένα συλλεγμένα από ΣμηΕΑ

3.3 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

3.3.1 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Η **Μηχανική Μάθηση** είναι μια εφαρμογή της τεχνητής νοημοσύνης (**AI**) που παρέχει στα συστήματα την δυνατότητα να μαθαίνουν και να βελτιώνονται αυτόματα από την εμπειρία, χωρίς να έχουν προγραμματιστεί ρητά. Η Μηχανική Μάθηση επικεντρώνεται στην ανάπτυξη προγραμμάτων ηλεκτρονικών υπολογιστών που έχουν πρόσβαση σε δεδομένα και μαθαίνουν από αυτά. Η διαδικασία μάθησης ξεκινά με παρατηρήσεις ή δεδομένα προκειμένου να αναζητήσουν μοτίβα και να λαμβάνουν καλύτερες αποφάσεις στο μέλλον με βάση τα παραδείγματα που τους έχουν δοθεί.

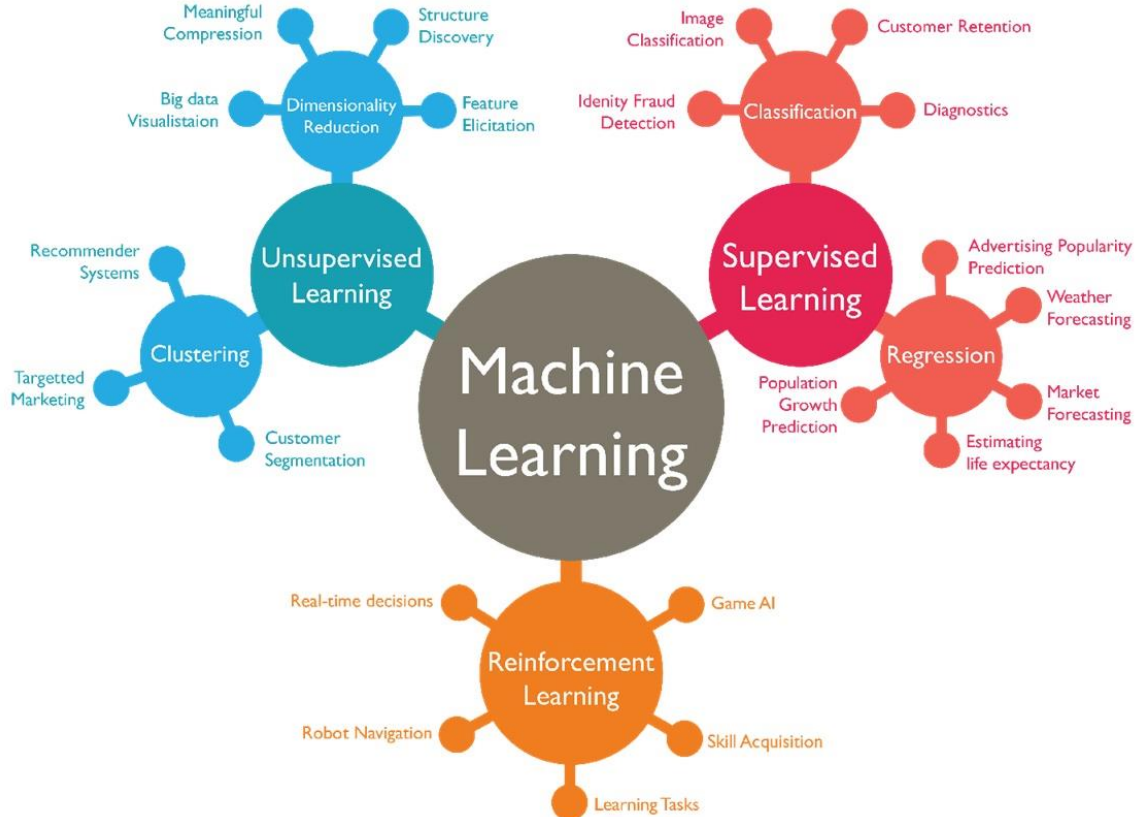
Υπάρχουν τρεις βασικές κατηγορίες αλγορίθμων μηχανικής μάθησης (Εικόνα 5):

- Οι **αλγόριθμοι με επίβλεψη** μπορούν να εφαρμόσουν ότι έχει μαθευτεί στο παρελθόν σε καινούργια δεδομένα χρησιμοποιώντας επισημασμένα παραδείγματα για να προβλέψουν μελλοντικά γεγονότα. Ξεκινώντας από την ανάλυση ενός γνωστού συνόλου δεδομένων κατάρτισης, ο αλγόριθμος παράγει μια συνάρτηση για να κάνει προβλέψεις σχετικά με τις τιμές εξόδου. Ο αλγόριθμος μπορεί επίσης να συγκρίνει την έξοδό του με την σωστή, προβλεπόμενη έξοδο ώστε να εντοπίσει λάθη και να τροποποιήσει το μοντέλο ανάλογα.
- Οι **αλγόριθμοι χωρίς επίβλεψη** εξετάζουν τον τρόπο με τον οποίο τα συστήματα μπορούν να καταλήξουν σε μία συνάρτηση για να περιγράψουν μια κρυφή δομή σε μη επισημασμένα δεδομένα. Το σύστημα δεν υπολογίζει την σωστή έξοδο, αλλά διερευνά τα δεδομένα και αντλεί συμπεράσματα για σύνολα δεδομένων.
- Οι **αλγόριθμοι ενισχυτικής μάθησης**, όπου ένας πράκτορας αλληλοεπιδρά με το περιβάλλον με σκοπό να μεγιστοποιήσει την μακροχρόνια επιβράβευση του. Η ενισχυτική μάθηση σχετίζεται με την θεωρία βέλτιστου ελέγχου.

3.3.2 ΕΝΙΣΧΥΤΙΚΗ ΜΑΘΗΣΗ

Η ενισχυτική μάθηση (**Reinforcement Learning**) είναι η επιστήμη της διαδοχικής λήψης αποφάσεων με τον βέλτιστο δυνατό τρόπο. Αποτελεί μια μέθοδο μηχανικής μάθησης όπου ένας πράκτορας παίρνει αποφάσεις με στόχο την μεγιστοποίηση μίας αθροιστικής ανταμοιβής. Διαφέρει από τις μεθόδους μηχανικής μάθησης με επίβλεψη, καθώς τα ζεύγη εισόδου/εξόδου δεν χρειάζεται να παρουσιαστούν και οι ενέργειες που δεν φέρνουν το επιθυμητό αποτέλεσμα δεν χρειάζεται να διορθωθούν ρητά. Ο πράκτορας δεν γνωρίζει την σωστή ενέργεια σε κάποια συγκεκριμένη χρονική στιγμή και προσπαθεί μέσω δοκιμών να την εντοπίσει.

Η βασική ιδέα της ενισχυτικής μάθησης είναι ότι υπάρχει ένα περιβάλλον που αντιπροσωπεύει τον εξωτερικό κόσμο ενός πράκτορα, ο οποίος ενεργεί πάνω σε αυτό το περιβάλλον και στην συνέχεια δέχεται παρατηρήσεις από αυτό, που περιλαμβάνουν μία ανταμοιβή και μια καινούργια κατάσταση (Εικόνα 6). Η ανταμοιβή είναι ο τρόπος με τον οποίο ο πράκτορας γνωρίζει πόσο καλή είναι η ενέργεια που επέλεξε.



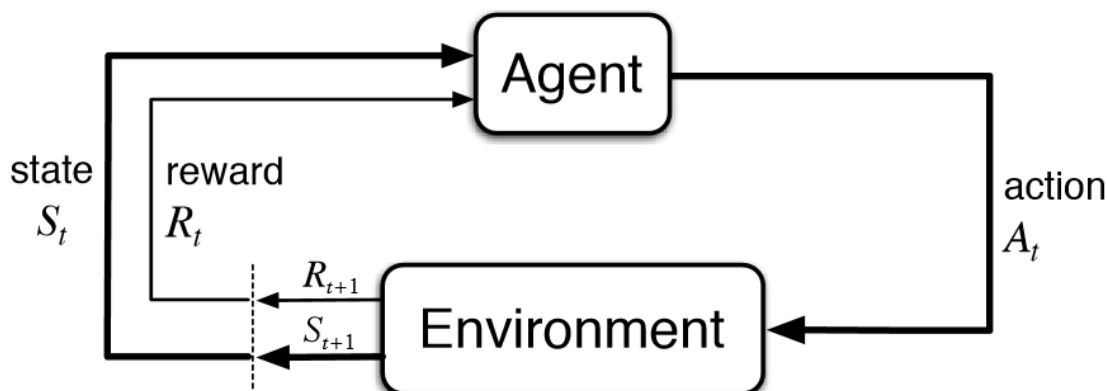
Εικόνα 5: Τύποι μηχανικής μάθησης (Πηγή: Shaikh 2017)

Τα περισσότερα προβλήματα ενισχυτικής μάθησης μπορούν να αναλυθούν σε ακολουθίες, στις οποίες ένας πράκτορας αλληλοεπιδρά με το περιβάλλον μέχρι να φτάσει σε μια τερματική κατάσταση που οδηγεί σε επαναφορά του συστήματος στην αρχική του κατάσταση. Κάθε ακολουθία αλληλεπίδρασης πράκτορα-περιβάλλοντος μεταξύ αρχικών και τερματικών καταστάσεων ονομάζεται **επεισόδιο**.

Η ενισχυτική μάθηση ορίζεται από τα ακόλουθα στοιχεία:

- **S**: το σύνολο των πιθανών καταστάσεων s του περιβάλλοντος
- **A**: το σύνολο των δυνατών ενεργειών a από τις οποίες μπορεί να επιλέξει ο πράκτορας σε κάθε χρονικό βήμα
- **$P(s'|s,a)$** : η κατανομή πιθανότητας μετάβασης. Για κάθε ζεύγος κατάσταση και ενέργειας, η συνάρτηση P εκφράζει την πιθανότητα το περιβάλλον να μεταβεί στην κατάσταση s'
- **$R(s,a,s')$** : η συνάρτηση ανταμοιβής
- **γ** : ο εκπτώτικος παράγοντας
- **ϵ** : ο βαθμός εξερεύνησης
- **l** : ο ρυθμός εκμάθησης

Το πρόβλημα της ενισχυτικής μάθησης ορίζεται ως προς τον βέλτιστο έλεγχο μιας διεργασίας Markov.



Εικόνα 6: Η Διαδικασία της ενισχυτικής μάθησης

Διεργασίες Markov

Η διεργασία Markov (**Markov Decision Process**) αποτελεί μια στοχαστική διαδικασία ελέγχου για διακριτές χρονικά διαδικασίες. Παρέχει ένα απλό πλαίσιο για την δόμηση του προβλήματος της ενισχυτικής μάθησης μέσω της αλληλεπίδρασης ενός πράκτορα με το περιβάλλον για την επίτευξη ενός στόχου. Ο πράκτορας αλληλοεπιδρά συνεχόμενα επιλέγοντας ενέργειες και το περιβάλλον απαντώντας σε αυτές τις ενέργειες και παρουσιάζοντας μια νέα κατάσταση στον πράκτορα και δίνοντας του μια ανταμοιβή.

Η όλη διεργασία βασίζεται πάνω στην **ιδιότητα Markov**, η οποία δηλώνει ότι **“Το μέλλον είναι ανεξάρτητο από το παρελθόν, δεδομένου του παρόντος”**. Μόλις γίνει γνωστή η παρούσα κατάσταση, το σύνολο των πληροφοριών που έχουν συναντηθεί μέχρι στιγμής μπορούν να αγνοηθούν και αυτή η κατάσταση αποτελεί επαρκές στατιστικό στοιχείο που δίνει τον ίδιο χαρακτηρισμό του μέλλοντος σαν να υπήρχε όλη η ιστορία διαθέσιμη. Σε μαθηματικούς όρους, μια κατάσταση S_t έχει την ιδιότητα Markov, αν και μόνο αν:

$$P[s_{t+1}] = P[s_{t+1} | s_1, s_2, \dots, s_t] \quad (1)$$

Η συνάρτηση πιθανότητας μετάβασης από μια Markov κατάσταση s σε μια κατάσταση διαδοχής s' είναι:

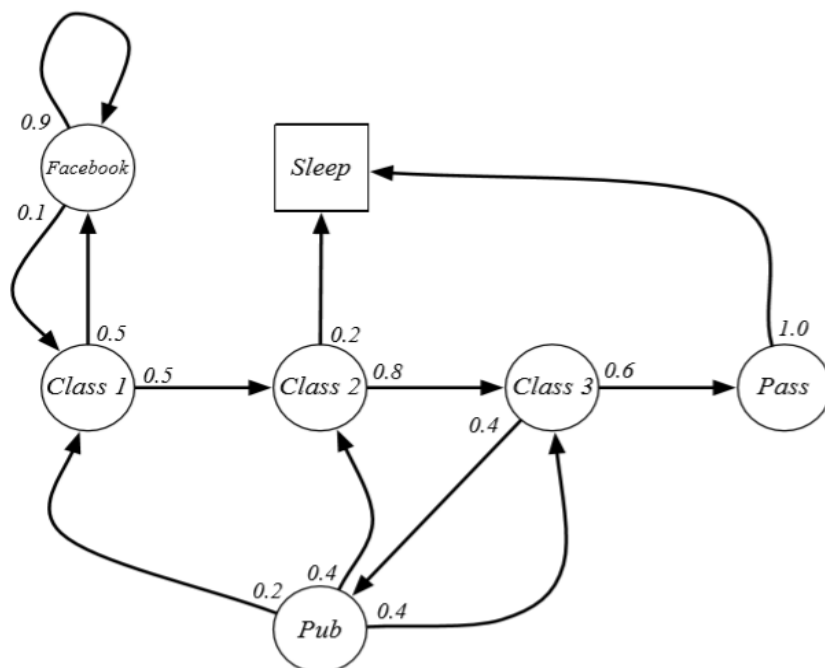
$$P_{ss'} = \mathbb{P}[s_{t+1} = s' | s_t = s] \quad (2)$$

και αποτελεί μια κατανομή πιθανότητας σε σχέση με τις επόμενες διαδοχικές καταστάσεις. Ο πίνακας μετάβασης P καθορίζει την πιθανότητα μετάβασης από όλες τις καταστάσεις s σε όλες τις επακόλουθες καταστάσεις s' .

$$P = \begin{matrix} & \text{to} \\ \text{from} & \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nm} \end{bmatrix} \end{matrix} \quad (3)$$

Όπου το άθροισμα κάθε γραμμής είναι 1.

Μια διεργασία Markov (ή αλυσίδα Markov) είναι μια τυχαία διαδικασία χωρίς μνήμη, δηλαδή μια ακολουθία από τυχαίες καταστάσεις (Εικόνα 7).



Εικόνα 7: Αλυσίδα Markov (Πηγή: Silver 2015)

Η Πολιτική

Η **πολιτική π** εκφράζει τον τρόπο με τον οποίο ο υπό εκπαίδευση πράκτορας συμπεριφέρεται σε κάθε χρονικό βήμα. Σε γενικές γραμμές, είναι η σύνδεση των αντιλαμβανόμενων καταστάσεων του περιβάλλοντος με τις αντίστοιχες βέλτιστες ενέργειες. Σε ορισμένες περιπτώσεις, η πολιτική είναι μια απλή συνάρτηση, ενώ σε άλλες μπορεί να περιλαμβάνει εκτεταμένες διαδικασίες όπως μια διαδικασία αναζήτησης. Η πολιτική **αποτελεί τον πυρήνα του πράκτορα** διότι από μόνη της είναι επαρκής για να καθορίσει την συμπεριφορά. Γενικά οι πολιτικές μπορεί να είναι στοχαστικές, καθορίζοντας πιθανότητες για κάθε ενέργεια.

Η πολιτική είναι μια κατανομή επί των ενεργειών δεδομένων των καταστάσεων της μορφής

$$\pi(a|s) = \mathbb{P}[a_t = a | s_t = s] \quad (4)$$

Η βέλτιστη πολιτική συμβολίζεται με π^* και αποτελεί τον ιδανικό τρόπο έτσι ώστε ο πράκτορας να αποκτήσει την μέγιστη δυνατή ανταμοιβή.

Συνάρτηση Ανταμοιβής

Σε κάθε χρονικό βήμα το περιβάλλον στέλνει στον πράκτορα έναν αριθμό ο οποίος αποτελεί την ανταμοιβή. Ο μοναδικός στόχος του πράκτορα είναι η μεγιστοποίηση της συνολικής ανταμοιβής που δέχεται μακροπρόθεσμα. Μέσω της **συνάρτησης ανταμοιβής προσδιορίζονται τα επιθυμητά και τα ανεπιθύμητα γεγονότα**. Το αντίστοιχο των ανταμοιβών σε ένα βιολογικό σύστημα είναι οι εμπειρίες της ευχαρίστησης και του πόνου. Είναι τα άμεσα και καθοριστικά χαρακτηριστικά του προβλήματος που αντιμετωπίζει ο πράκτορας. Η συνάρτηση ανταμοιβής είναι της μορφής

$$R_S = \mathbb{E}[r_{t+1} | s_t = s] \quad (5)$$

Το **σήμα ανταμοιβής** αποτελεί την κύρια βάση για την τροποποίηση της πολιτικής που ακολουθεί ο πράκτορας· εάν μια ενέργεια που επιλέγεται από την πολιτική ακολουθείται από χαμηλή ανταμοιβή, τότε η πολιτική μπορεί να αλλάξει έτσι ώστε σε περίπτωση που ξαναβρεθεί στην ίδια κατάσταση να επιλέξει μια διαφορετική ενέργεια. Γενικά, τα μηνύματα ανταμοιβής μπορεί να είναι στοχαστικές συναρτήσεις της κατάστασης του περιβάλλοντος και των δράσεων που επιλέγονται

Το άθροισμα των ανταμοιβών ονομάζεται **επιστροφή**, G_t , και προσδιορίζεται ως,

$$G_t = r_t + r_{t+1} + r_{t+2} + \dots + r_{N-1} \quad (6)$$

όπου το N υποδηλώνει το τέλος ενός επεισοδίου και το t τον χρονικό δείκτη.

Για συνεχιζόμενες εργασίες ορίζεται η εκπτώτικη επιστροφή ως,

$$G_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k} \quad (7)$$

όπου $\gamma \in [0, 1)$ ο εκπτώτικος παράγοντας, ο οποίος έχει την ικανότητα να χαρακτηρίσει την σημασία των ανταμοιβών που λήφθηκαν νωρίτερα σε σχέση με εκείνες που λήφθηκαν αργότερα. Όσο μεγαλύτερος είναι ο εκπτώτικος παράγοντας τόσο περισσότερο σημαντικές θεωρούνται οι μελλοντικές ανταμοιβές στο σύστημα μας, ενώ μικρός παράγοντας σημαίνει ότι μόνο οι τωρινές ανταμοιβές λαμβάνονται υπόψιν. Η ύπαρξη του γ εξυπηρετεί την αποφυγή άπειρων κυκλικών διεργασιών.

Συνάρτηση Αξίας

Ενώ η ανταμοιβή υποδεικνύει τι είναι καλό βραχυπρόθεσμα, η **συνάρτηση αξίας προσδιορίζει τι είναι καλό μακροπρόθεσμα**. Σε γενικές γραμμές, η αξία μιας κατάστασης είναι η συνολική ανταμοιβή που ένας πράκτορας αναμένει να λάβει στο μέλλον, ξεκινώντας από αυτή την κατάσταση. Ενώ οι ανταμοιβές καθορίζουν το άμεσο όφελος που μπορεί να επιφέρει μια κατάσταση, οι αξίες υποδεικνύουν την μακροπρόθεσμη ελκυστικότητα καταστάσεων αφού ληφθούν υπόψιν οι πιθανές καταστάσεις που θα ακολουθήσουν και τις διαθέσιμες ανταμοιβές αυτών. Για παράδειγμα, μια κατάσταση μπορεί πάντα να επιφέρει μικρή άμεση ανταμοιβή αλλά να έχει υψηλή αξία επειδή συχνά ακολουθείται από καταστάσεις που επιφέρουν μεγάλες ανταμοιβές.

Οι ανταμοιβές είναι κατά μία έννοια πρωταρχικές, ενώ **οι αξίες**, ως προβλέψεις των ανταμοιβών, **είναι δευτερεύουσες**. Χωρίς ανταμοιβές δεν θα υπήρχαν αξίες, ο μόνος λόγος που εκτιμώνται οι αξίες είναι για να επιτευχθούν μεγαλύτερες ανταμοιβές. Παρόλα αυτά, οι αξίες είναι εκείνες που λαμβάνονται, κατά κύριο λόγο, υπόψιν στην αξιολόγηση των αποφάσεων. Οι επιλογές ενεργειών βασίζονται στις εκτιμήσεις των αξιών. Επιλέγονται ενέργειες που οδηγούν σε καταστάσεις με μεγαλύτερες αξίες και όχι μεγαλύτερες ανταμοιβές, διότι αυτές οι ενέργειες θα επιφέρουν μεγαλύτερες ανταμοιβές μακροπρόθεσμα. Αυτό στην πράξη είναι αρκετά δύσκολο, διότι οι ανταμοιβές δίνονται απευθείας από το περιβάλλον, ενώ οι αξίες πρέπει να υπολογίζονται και να διορθώνονται ανάλογα με τις παρατηρήσεις του πράκτορα κατά την διάρκεια της ζωής του. Προκειμένου να είναι σε θέση να αποφασίσει ποια ενέργεια θα επιλέξει σε μια συγκεκριμένη στιγμή, είναι σημαντικό για τον πράκτορα να γνωρίζει πόσο καλή είναι η κάθε κατάσταση. Γι' αυτό και το **πιο σημαντικό στοιχείο** όλων σχεδόν των αλγορίθμων ενισχυτικής μάθησης **είναι η μέθοδος με την οποία υπολογίζονται οι αξίες**. Η συνάρτηση αξίας για την πολιτική π δίνεται από

$$V_{\pi}(s) = \mathbb{E}_{\pi}(G_t | s_t = s) = \mathbb{E}_{\pi}(\sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s) \quad (8)$$

Ομοίως, η **συνάρτηση ενέργειας-αξίας**, γνωστή και ως Q-συνάρτηση, ορίζεται ως το αναμενόμενο άθροισμα των ανταμοιβών ενώ επιλέγεται η ενέργεια a στην κατάσταση s και μετέπειτα ακολουθείται η πολιτική π. Η συνάρτηση δίνεται από

$$Q_{\pi}(s, a) = \mathbb{E}_{\pi}(\sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s, a_t = a) \quad (9)$$

3.3.3 ΟΙ ΕΞΙΣΩΣΕΙΣ BELLMAN

Οι εξισώσεις Bellman διατυπώνουν το πρόβλημα της μεγιστοποίησης των αναμενόμενων αθροιστικών ανταμοιβών από την άποψη μιας συνάρτησης αξίας. Μια πολιτική π θεωρείται καλύτερη από μια πολιτική π' αν η εκτιμώμενη επιστροφή αυτής της πολιτικής είναι μεγαλύτερη από την αντίστοιχη της π' για κάθε $s \in S$, που συνεπάγεται, $V^{\pi}(s) \geq V^{\pi'}(s)$ για κάθε $s \in S$. Συνεπώς η βέλτιστη συνάρτηση αξίας $V^*(s)$ είναι,

$$V^*(s) = \max_{\pi} V_{\pi}(s), \quad \forall s \in S \quad (10)$$

Ομοίως, η βέλτιστη συνάρτηση ενέργειας-αξίας $Q^*(s, a)$ μπορεί να οριστεί ως,

$$Q^*(s, a) = \max_{\pi} Q_{\pi}(s, a), \quad \forall s \in S, a \in A \quad (11)$$

Επίσης για μια βέλτιστη πολιτική μπορεί να γραφτεί η εξής εξίσωση,

$$V^*(s) = \max_{a \in A(s)} Q_{\pi^*}(s, a) \quad (12)$$

Αντικαθιστώντας σε αυτήν την συνάρτηση ενέργειας-αξίας,

$$V^*(s) = \max_a \mathbb{E}_{\pi^*}(G_t | s_t = s, a_t = a)$$

$$\begin{aligned}
&= \max_a \mathbb{E}_\pi \left(\sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t = s, a_t = a \right) \\
&= \max_a \sum_{s'} p(s' \mid s, a) [R(s, a, s') + \gamma \max_{a'} Q^*(s', a')] \quad (13)
\end{aligned}$$

Αυτή η εξίσωση είναι γνωστή ως **η εξίσωση βελτιστοποίησης του Bellman** για την $V^*(s)$. Η εξίσωση βελτιστοποίησης του Bellman για την Q^* είναι

$$\begin{aligned}
Q^*(s, a) &= \mathbb{E}(r_t + \gamma \max_{a'} Q^*(s_{t+1}, a') \mid s_t = s, a_t = a) \\
&= \sum_{s'} p(s' \mid s, a) [R(s, a, s') + \gamma \max_{a'} Q^*(s', a')] \quad (14)
\end{aligned}$$

Εάν οι πιθανότητες μετάβασης και οι συναρτήσεις ανταμοιβής είναι γνωστές, οι εξισώσεις βελτιστοποίησης του Bellman μπορούν να επιλυθούν με επαναληπτικό τρόπο. Αυτή η προσέγγιση είναι γνωστή ως **Δυναμικός Προγραμματισμός**. Οι αλγόριθμοι που υποθέτουν ότι αυτές οι πιθανότητες είναι γνωστές ή εκτιμώνται είναι συλλογικά γνωστοί ως **αλγόριθμοι με βάση μοντέλα**.

Αλλά οι περισσότεροι αλγόριθμοι, υποθέτουν ότι οι πιθανότητες αυτές δεν είναι γνωστές και εκτιμούν την πολιτική και την συνάρτηση αξίας πραγματοποιώντας επαναλήψεις στο σύστημα. Αυτές οι μέθοδοι είναι γνωστές ως **αλγόριθμοι χωρίς μοντέλο**. Η μέθοδος Monte Carlo, οι μέθοδοι χρονικής διαφοροποίησης και αναζήτησης πολιτικής είναι οι πιο συνηθισμένοι αλγόριθμοι χωρίς μοντέλο.

3.3.4 ΜΕΘΟΔΟΙ ΧΩΡΙΣ ΜΟΝΤΕΛΟ

Οι μέθοδοι χωρίς μοντέλο μπορούν να εφαρμοστούν σε οποιοδήποτε πρόβλημα ενισχυτικής μάθησης, καθώς δεν απαιτούν κάποιο μοντέλο του περιβάλλοντος. Οι περισσότερες προσεγγίσεις χωρίς μοντέλο είτε προσπαθούν να μάθουν μια συνάρτηση αξίας και να καταλήξουν σε μια βέλτιστη πολιτική από αυτήν (**μέθοδοι βασισμένες στην συνάρτηση αξίας**), είτε προσπαθούν απευθείας να αναπτύξουν μια βέλτιστη πολιτική αναζητώντας τις κατάλληλες παραμέτρους (**μέθοδοι αναζήτησης πολιτικής**).

Οι προσεγγίσεις χωρίς μοντέλο μπορούν επίσης να ταξινομηθούν ως **εντός πολιτικής ή εκτός-πολιτικής (on-policy or off-policy)**. Οι μέθοδοι εντός-πολιτικής χρησιμοποιούν την τρέχουσα πολιτική για να επιλέξουν ενέργειες και με βάση αυτές επικαιροποιούν την τρέχουσα πολιτική. Για παράδειγμα, μια μέθοδος θεωρείται εντός-πολιτικής αν ενημερώνει τις Q-αξίες χρησιμοποιώντας την Q-αξία της επόμενης κατάστασης s' και της ενέργειας της τωρινής πολιτικής a , δηλαδή υπολογίζει την επιστροφή ζευγών κατάστασης-ενέργειας υποθέτοντας ότι η τρέχουσα πολιτική θα συνεχίσει να ακολουθείται. Οι μέθοδοι εκτός-πολιτικής χρησιμοποιούν μια διαφορετική διερευνητική πολιτική για να επιλέξουν ενέργειες, καθώς ένα μοντέλο εκτός-πολιτικής μαθαίνει την αξία της βέλτιστης πολιτικής ανεξάρτητα από τις ενέργειες του πράκτορα. Η πολιτική που χρησιμοποιείται για να διαμορφώσει την συμπεριφορά, που ονομάζεται πολιτική συμπεριφοράς, μπορεί να μην συσχετίζεται με την πολιτική που αξιολογείται και βελτιώνεται, που ονομάζεται εκτιμώμενη πολιτική. Η **Q-Μάθηση** αποτελεί μια εκτός-πολιτικής μέθοδο.

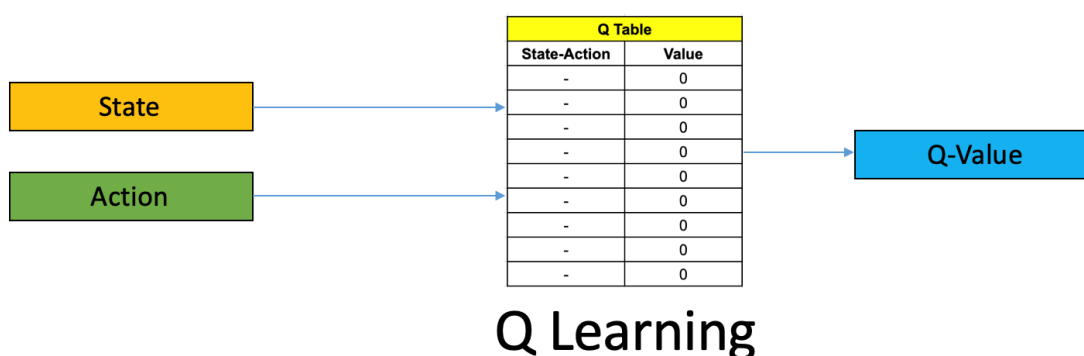
Η Q-Μάθηση αποτελεί έναν αλγόριθμο ενισχυτικής μάθησης χωρίς μοντέλο (Εικόνα 8). Ο σκοπός της μεθόδου είναι να μάθει μια πολιτική, που επιλέγει την κατάλληλη ενέργεια που

πρέπει να κάνει ο πράκτορας για τις διάφορες συνθήκες. Δεν απαιτεί μοντέλο του περιβάλλοντος και μπορεί να διαχειριστεί προβλήματα με στοχαστικές μεταβάσεις και ανταμοιβές, χωρίς την ανάγκη για προσαρμογές. Η σχέση ενημέρωσης των Q-αξιών είναι:

$$Q(s, a) = (1 - lr)Q(s, a) + lr[r_t + \gamma Q^*(s', a')] \quad (15)$$

όπου, $Q^*(s', a')$ είναι η μέγιστη αξία της επόμενης κατάστασης.

Για κάθε πεπερασμένη Διεργασία Markov (Finite Markov Decision Process), η μέθοδος βρίσκει μια βέλτιστη πολιτική δεδομένου του άπειρου χρόνου εξερεύνησης και μιας εν μέρει τυχαίας πολιτικής. Το Q είναι το όνομα της συνάρτησης που επιστρέφει την ανταμοιβή και μπορεί να θεωρηθεί ότι αντιπροσωπεύει την “ποιότητα” (**Quality**) μιας ενέργειας που έχει ληφθεί σε μια δεδομένη κατάσταση.



Εικόνα 8: Ο αλγόριθμος Q-Μάθησης (Πηγή: Choudhary 2019)

Η Q-Μάθηση αποτελεί μια εκτός-πολιτικής μέθοδο διότι ενημερώνει τις Q-αξίες χρησιμοποιώντας την τιμή της Q-αξίας της επόμενης κατάστασης s' και της άπληστης ενέργειας a' (εξηγείται αργότερα). Με άλλα λόγια, υπολογίζει την επιστροφή (αθροιστική εκπτώτικη μελλοντική ανταμοιβή) για ζεύγη κατάστασης-ενέργειας, υποθέτοντας ότι ακολουθείται μια άπληστη πολιτική παρόλο που στην πραγματικότητα δεν ακολουθείται άπληστη πολιτική.

3.3.5 ΕΞΕΡΕΥΝΗΣΗ ΕΝΑΝΤΙ ΕΚΜΕΤΑΛΛΕΥΣΗΣ

Εξερεύνηση είναι η συλλογή επιπλέον πληροφοριών προκειμένου να οδηγηθεί το μοντέλο στην λήψη καλύτερων αποφάσεων στο μέλλον, ενώ εκμετάλλευση είναι η επιλογή των βέλτιστων αποφάσεων με βάση την τωρινή γνώση.

Ο κάθε συνδυασμός κατάστασης-ενέργειας έχει μια αξία. Αν ήταν γνωστές οι πραγματικές αξίες της κάθε ενέργειας για κάθε κατάσταση τότε για κάθε κατάσταση θα υπήρχε μια γνωστή βέλτιστη ενέργεια, η οποία και θα επιλεγόταν. Ωστόσο στην πραγματικότητα δεν είναι γνωστές αυτές οι αξίες και όσο το μοντέλο εκτίθεται σε νέα δεδομένα, οι αξίες αυτές ανανεώνονται και αποτελούν εκτιμήσεις των πραγματικών αξιών. Έτσι σε κάθε χρονική στιγμή και για κάθε κατάσταση υπάρχει τουλάχιστον μια ενέργεια της οποίας η εκτιμώμενη αξία είναι η μεγαλύτερη, αυτή ονομάζεται **άπληστη ενέργεια**. Όταν επιλέγεται μία από αυτές τις ενέργειες, γίνεται εκμετάλλευση της μέχρι τώρα γνώσης. Αντιθέτως αν επιλεγεί

μία από τις μη άπληστες ενέργειες, τότε γίνεται εξερεύνηση, διότι αυτό βοηθάει στην βελτιωμένη εκτίμηση των αξιών των μη άπληστων ενεργειών.

Η εκμετάλλευση είναι η σωστή τακτική για την μεγιστοποίηση των αναμενόμενων ανταμοιβών σε κάποιο χρονικό βήμα, ωστόσο η εξερεύνηση μπορεί να οδηγήσει σε μεγαλύτερες αθροιστικές ανταμοιβές μακροπρόθεσμα.

Ένα σημαντικό κομμάτι στην ανάπτυξη ενός μοντέλου ενισχυτικής μάθησης είναι η **εξισορρόπηση μεταξύ εξερεύνησης και εκμετάλλευσης**. Η παράμετρος ϵ δείχνει τι ποσοστό του χρόνου θα εξερευνώνται ενέργειες οι οποίες δεν έχουν δοκιμαστεί μέχρι στιγμής και θα αγνοείται η μέχρι τώρα γνώση για την επιλογή της βέλτιστης ενέργειας.

3.3.6 ΡΥΘΜΟΣ ΕΚΜΑΘΗΣΗΣ

Ο ρυθμός εκμάθησης καθορίζει σε **ποιο βαθμό οι καινούργιες πληροφορίες υπερισχύουν των παλιών**. Μια τιμή της παραμέτρου ίση με 0 οδηγεί τον πράκτορα στο να μην μαθαίνει τίποτα (αποκλειστική εκμετάλλευση προηγούμενης γνώσης), ενώ μια τιμή ίση με 1 οδηγεί τον πράκτορα στο να λαμβάνει υπόψιν μόνο τις καινούργιες πληροφορίες αγνοώντας προϋπάρχουσα γνώση.

Σε πλήρως ντετερμινιστικά περιβάλλοντα, ρυθμός εκμάθησης ίσος με 1 είναι ιδανικός. Σε στοχαστικά προβλήματα, ο αλγόριθμος συγκλίνει κάτω από ορισμένες τεχνικές συνθήκες στο ρυθμό εκμάθησης που πρέπει να μειωθεί σταδιακά στο μηδέν. Στην πράξη, επιλέγεται σταθερός ρυθμός εκμάθησης ίσος με 0.1 για παράδειγμα.

4. ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

4.1 ΕΙΣΑΓΩΓΗ

Η χρήση **Συστημάτων μη Επανδρωμένων Αεροσκαφών** (ΣμηΕΑ) έχει προταθεί ως ένα αποτελεσματικό και πρωτοποριακό εργαλείο των ευφυών συστημάτων μεταφορών (Barmpounakis et al. 2016), για την παρακολούθηση της κυκλοφορίας σε μεγάλα δίκτυα όπου δεν μπορούν να καλυφθούν από σταθερές κάμερες καθώς και για την ανάλυση διαφόρων κυκλοφοριακών φαινομένων. Τα ΣμηΕΑ αποτελούν μια νέα και οικονομικά αποδοτική λύση για την συλλογή μαζικών δεδομένων τροχιών οχημάτων από οδικές αρτηρίες και αντικαθιστούν την παλιά προσέγγιση η οποία περιελάμβανε τη χρήση ήδη εγκατεστημένων καμερών.

Η χρήση ΣμηΕΑ σε αστικές περιοχές αποτελεί έναν **ιδανικό τρόπο παρακολούθησης και ανάλυσης της κυκλοφορίας** και έχει κινήσει το ενδιαφέρον αρκετών συγκοινωνιακών ερευνητών. Ωστόσο, τα υπάρχοντα πειράματα σε όλο τον κόσμο πραγματοποιούνται σε πολύ μικρή κλίμακα, συνήθως με ένα drone που καλύπτει μια ή δύο διασταυρώσεις ή ένα τμήμα οδού.

4.2 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα για την συγκεκριμένη διπλωματική εργασία λήφθηκαν από το πείραμα που πραγματοποιήθηκε στο κέντρο της Αθήνας με χρήση Συστημάτων μη Επανδρωμένων Αεροσκαφών (drones) (Barmpounakis et al. 2019). Σκοπός του πειράματος ήταν η **καταγραφή κυκλοφοριακών ροών σε αστικό περιβάλλον** και η παροχή μιας ολικής εικόνας για το πώς τα ιδιαίτερα χαρακτηριστικά των ΣμηΕΑ μπορούν να ξεπεράσουν υφιστάμενους περιορισμούς στην καταγραφή της κυκλοφορίας.

Το πείραμα αποτέλεσε το πρώτο μεγάλης κλίμακας του είδους του με στόχο την καταγραφή ροών κυκλοφορίας σε αστικό περιβάλλον. Πιο συγκεκριμένα, χρησιμοποιήθηκε **στόλος 10 ΣμηΕΑ σε μία ευρύτερη περιοχή του κέντρου της Αθήνας**. Επιλέχθηκε ο Κεντρικός Τομέας Αθηνών, καθώς αποτελεί ένα αστικό και πολυάσχολο περιβάλλον, στο οποίο πραγματοποιούνται μετακινήσεις με διάφορα μέσα και που μπορεί να επιτρέψει την παρακολούθηση διαφόρων κυκλοφοριακών φαινομένων. Το πείραμα πραγματοποιήθηκε κατά την διάρκεια της πρωινής αιχμής (8:00-10:30), καθώς με βάση την ισχύουσα νομοθεσία δεν επιτρέπεται η χρήση ΣμηΕΑ κατά τις νυχτερινές ώρες, για κάθε εργάσιμη ημέρα της εβδομάδας. Η συλλογή των δεδομένων, λόγω των περιορισμών των ΣμηΕΑ, έγινε σε φάσεις διάρκειας 30 λεπτών, διότι η αυτονομία των ΣμηΕΑ είναι 25 λεπτά.

Η τοποθέτηση των 10 ΣμηΕΑ καθώς και τα δύο σημεία ελέγχου, απογείωσης και προσγείωσής τους εμφανίζεται στην Εικόνα 9. Τα **δεδομένα** που χρησιμοποιήθηκαν στην εκπαίδευση του μοντέλου της εργασίας **πάρθηκαν από τη Λεωφόρο Αλεξάνδρας** (σημείο 8). Η Λεωφόρος Αλεξάνδρας επιλέχθηκε ως μια κεντρική οδός με υψηλούς κυκλοφοριακούς φόρτους και πολλές λωρίδες κυκλοφορίας ανά κατεύθυνση (3

λωρίδες/κατεύθυνση). Έτσι δίνεται η δυνατότητα στο μοντέλο να εκτεθεί σε μια πληθώρα καταστάσεων και να προσαρμοσθεί στην καλύτερη δυνατή διαχείρισή τους.



Εικόνα 9: Σημεία τοποθέτησης ΣμηΕΑ (Πηγή: Barmounakis et al. 2019)

Η χρήση ΣμηΕΑ μπορεί να αποτελέσει έναν αποτελεσματικό και οικονομικό τρόπο για την **ανάλυση μικροσκοπικών χαρακτηριστικών της κυκλοφορίας**, και με σωστή βαθμονόμηση και σταθεροποίηση του βίντεο μπορούν να εξαχθούν αποτελέσματα μεγάλης ακρίβειας με σφάλμα ταχυτήτων που δεν ξεπερνά τα 1.2 km/h (Barmounakis et al. 2017). Επίσης η ταχύτητα του εξεταζόμενου οχήματος και το υψόμετρο του αεροσκάφους δεν επηρεάζουν την ακρίβεια των δεδομένων.

Τα βίντεο καταγράφηκαν σε ανάλυση 4K (4096x2160) στα 25 καρέ ανά δευτερόλεπτο από DJI Drones. Δεδομένης της μεγάλης εξέλιξης στον τομέα της μηχανικής όρασης και της ανάγκης για εξασφάλιση υψηλού επιπέδου ακρίβειας, **η ανάλυση των βίντεο και η εξόρυξη των τροχιών** των οχημάτων έγινε από μια νεοσύστατη εταιρεία που εξειδικεύεται στην εξόρυξη δεδομένων, την **DataFromSky**. Η ακρίβεια της εξαγωγής αγγίζει το 99.8% (Barmounakis et al. 2019), ενώ έγινε ανθρώπινη επέμβαση για περαιτέρω διόρθωση λαθών και για αντιμετώπιση των ζητημάτων που σχετίζονται με την προοπτική άποψη των ΣμηΕΑ.

Τα αποτελέσματα της ανάλυσης περιλαμβάνουν λεπτομερείς τροχιές των εξεταζόμενων οχημάτων βαθμονομημένες στο **σύστημα συντεταγμένων WGS-84**. Η χρονική συχνότητα είναι 0.04 sec, η μεγαλύτερη που επέτρεπε η κάμερα. Εκτός από τις συντεταγμένες, έγινε υπολογισμός και της ταχύτητας (πρώτη παράγωγος της θέσης), της διαμήκης και της πλευρικής επιτάχυνσης (δεύτερη παράγωγος της θέσης), της συνολικής

διανυθείσας απόστασης, της μέσης ταχύτητας και του τύπου οχήματος. Οι κατηγορίες των οχημάτων είναι:

1. Ποδήλατο
2. Δίκυκλο
3. Αυτοκίνητο
4. Ταξί
5. Ημιφορτηγό
6. Φορτηγό
7. Λεωφορείο

Η αναπαράσταση των διαφόρων κινήσεων σε μία εκτεταμένη περιοχή του πειράματος παρουσιάζεται στην Εικόνα 10.



Εικόνα 10: Αναπαράσταση των διαφόρων κινήσεων σε μια εκτεταμένη περιοχή (Πηγή: *Bartrounakis et al. 2019*)

4.3 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα ήταν αρχικά διαχωρισμένα σε 5 αρχεία csv, ένα για κάθε χρονική περίοδο. Το κάθε αρχείο αποτελείται από γραμμές, όσες και τα οχήματα που εμφανίστηκαν κατά την διάρκεια της εξεταζόμενης περιόδου, στο οδικό τμήμα που κάλυπτε το ΣμηΕΑ. Σε κάθε γραμμή περιέχονται όλες οι πληροφορίες που συγκεντρώθηκαν για το κάθε όχημα την κάθε χρονική στιγμή (**συντεταγμένες, ταχύτητα, επιταχύνσεις**). Η συνολική απόσταση δίνεται σε μέτρα, η μέση ταχύτητα σε χλμ/ώρα, οι συντεταγμένες σε βαθμούς, η ταχύτητα σε χλμ/ώρα, οι επιταχύνσεις σε μέτρα/δευτερόλεπτο² ενώ ο χρόνος σε χιλιοστά του δευτερολέπτου.

Η μορφή του πίνακα εμφανίζεται παρακάτω, με τις στήλες Trajectory lat, lon , Speed, Tan. Accel, Lat. Accel, Time να επαναλαμβάνονται για κάθε χρονική στιγμή (Πίνακας 1).

Πίνακας 1: Δομή των δεδομένων τροχιών

Track ID	Type	Traveled Dist. [m]	Avg. Speed [km/h]	Trajectory[lat [deg]	lon [deg]	Speed [km/h]	Tan. Accel. [ms-2]	Lat. Accel. [ms-2]	Time [ms]
1	Medium Vehicle	375.28	23.2130	37.9914	23.7323	11.1338	-0.0779	0.0523	12000
2	Car	518.37	9.5308	37.9885	23.7291	32.9903	0.3480	-0.0038	12200
3	Car	157.22	36.7533	37.9902	23.7293	31.0879	-0.1004	0.0188	12200
4	Heavy Vehicle	350.69	31.8811	37.9886	23.7291	29.2202	-0.1182	-0.1445	12200
5	Car	497.14	23.4255	37.9910	23.7364	55.6260	0.1025	0.0104	12400
6	Car	427.66	8.2330	37.9902	23.7293	31.3172	-0.0307	0.0133	14000
7	Car	850.43	12.5886	37.9885	23.7291	34.9722	0.3087	-0.0703	14200
8	Car	831.82	19.2699	37.9910	23.7364	61.1052	0.0092	-0.0262	14400
9	Car	329.05	13.1328	37.9886	23.7291	28.3907	-0.2543	-0.0576	14600
10	Car	847.80	12.8347	37.9886	23.7291	40.6251	0.3589	-0.0250	14600
11	Motorcycle	236.79	27.6768	37.9886	23.7291	45.5406	0.0332	0.0230	14800
12	Car	636.95	24.5505	37.9910	23.7363	54.8595	0.1071	0.0134	15600
13	Car	241.80	6.6857	37.9885	23.7291	35.2763	0.1398	0.0001	16200
14	Car	249.04	7.4587	37.9886	23.7291	36.8343	0.1524	-0.0402	16200
15	Car	43.74	8.9461	37.9923	23.7301	21.7403	0.1399	0.0861	17000
16	Car	483.95	26.3971	37.9886	23.7291	37.7706	-0.4250	-0.1424	17800
17	Car	637.47	23.3220	37.9910	23.7364	56.2176	0.1164	0.0139	18400
18	Motorcycle	41.59	8.3183	37.9923	23.7301	14.4228	-0.3637	0.1845	18400
19	Car	239.47	23.5542	37.9885	23.7291	36.5177	0.6641	-0.0284	18600
20	Heavy Vehicle	495.71	23.6052	37.9911	23.7363	56.2237	0.1540	0.0115	18800
21	Car	415.15	31.6640	37.9910	23.7363	56.2337	0.1321	0.0025	20600
22	Car	830.54	21.6349	37.9910	23.7363	63.9345	0.2450	0.0010	21800
23	Bus	634.09	17.5324	37.9910	23.7363	51.2941	0.1534	-0.0093	22000
24	Motorcycle	244.34	8.1297	37.9885	23.7291	33.4819	0.2645	0.0277	22200
25	Heavy Vehicle	344.33	20.8684	37.9886	23.7291	29.5654	-0.2496	-0.2222	22200
26	Motorcycle	472.76	32.8562	37.9926	23.7314	34.0226	-0.0690	-2.0088	22400
27	Motorcycle	541.73	38.2398	37.9926	23.7315	33.1216	0.1999	0.0081	22600
28	Car	494.06	19.9843	37.9910	23.7363	52.6961	0.0428	0.0281	24600
29	Motorcycle	541.61	39.1523	37.9925	23.7315	38.4969	0.5359	-0.0065	24800
30	Medium Vehicle	242.92	4.4940	37.9885	23.7291	29.0284	0.4372	-0.0079	25000

Προκαταρκτική επεξεργασία

Η προκαταρκτική επεξεργασία των δεδομένων έγινε με χρήση της γλώσσας R στο περιβάλλον του R-Studio, όπου οριοθετήθηκε ο αριθμός των στηλών του κάθε αρχείου στις 50,000 για καλύτερη επεξεργασία και μείωση του υπολογιστικού φόρτου και έγινε αφαίρεση της γραμμής που περιείχε τα ονόματα των στηλών έτσι ώστε να γίνει σωστά η ανάγνωση του αρχείου csv από την rython στην συνέχεια.

Η επεξεργασία των δεδομένων έγινε κατά κύριο λόγο με χρήση της γλώσσας προγραμματισμού rython στο περιβάλλον Spyder που αποτελεί μέρος του Anaconda. Η παρακάτω διαδικασία ακολουθήθηκε για κάθε ένα από τα 5 αρχεία ξεχωριστά. Αρχικά δημιουργήθηκε ένας **τρισδιάστατος πίνακας, όπου η μία διάσταση είναι ο χρόνος**. Έτσι σε κάθε χρονική στιγμή αντιστοιχεί ένας δυσδιάστατος πίνακας ο οποίος περιέχει στις γραμμές τα οχήματα που εμφανίστηκαν στην οδό την συγκεκριμένη χρονική στιγμή και στις στήλες τα δεδομένα του κάθε οχήματος. Με αυτόν τον τρόπο η διαχείριση των δεδομένων ήταν αρκετά πιο εύκολη στα επόμενα βήματα, καθώς τα δεδομένα ήταν καλά διαχωρισμένα.

Επειδή η συχνότητα των μετρήσεων ήταν πολύ μεγάλη, ανά 0.04 sec, **οι μετρήσεις που κρατήθηκαν ήταν ανά 0.8 sec**, διότι οι μετρήσεις με μεγάλη συχνότητα έχουν πολύ μεγάλη συσχέτιση μεταξύ τους. Έτσι η κατάσταση της κυκλοφορίας σε δύο διαδοχικές χρονικές στιγμές διαφοράς 0.04 sec είναι πολύ κοντινή, με τα οχήματα να μην έχουν μετατοπιστεί ιδιαίτερα. Αυτό σημαίνει ότι αν λαμβάνονταν υπόψιν όλες οι χρονικές στιγμές θα υπήρχε πολύ μεγάλος υπολογιστικός φόρτος χωρίς να δίνεται η δυνατότητα στο μοντέλο να προσαρμοσθεί σε διαφορετική κατάσταση της κυκλοφορίας.

Μετατροπή των συντεταγμένων

Στην συνέχεια έγινε **μετατροπή των συντεταγμένων από το σύστημα WGS 84 στο προβολικό σύστημα UTM**. Το Παγκόσμιο Γεωκεντρικό Σύστημα (World Geodetic System) 84 αποτελεί το πρότυπο σύστημα συντεταγμένων παγκοσμίως, χρησιμοποιείται από το δορυφορικό σύστημα GPS και περιγράφει την γη ως ελλειψοειδές. Ωστόσο το ελλειψοειδές δεν είναι αναπτυσσόμενη επιφάνεια, με αποτέλεσμα να είναι αδύνατη η απεικόνιση του στο επίπεδο. Για τον λόγο αυτό γίνεται Εγκάρσια Μερκατορική Προβολή (UTM) των συντεταγμένων. Το σύστημα αυτό χωρίζει τη γη σε 60 ζώνες, γεωγραφικού μήκους 6°, με την πρώτη στον μεσημβρινό του Greenwich. Εφόσον όλα τα εξεταζόμενα σημεία βρίσκονται εντός μίας ζώνης, οι παραμορφώσεις είναι αρκετά μικρές (<1:2500). Για να αποφευχθούν αρνητικές συντεταγμένες η τιμή των τετμημένων των κεντρικών μεσημβρινών είναι 500000 m. Η αρχή των τεταγμένων είναι η τομή του κεντρικού μεσημβρινού με τον ισημερινό. Η εξεταζόμενη περιοχή της Αθήνας βρίσκεται στην ζώνη 34.

Αφαίρεση οχημάτων εκτός περιοχής μελέτης

Αφού γίνει μετατροπή των συντεταγμένων, γίνεται **αφαίρεση όσων οχημάτων δεν βρίσκονται επί της Λεωφόρου Αλεξάνδρας**, με βάση τις συντεταγμένες τους, και γίνεται χαρακτηρισμός του κάθε οχήματος ανάλογα με τον τύπο του. Προκύπτουν οι εξής κατηγορίες:

- Δίκυκλο (1)
- Αυτοκίνητο (2)
- Ταξί (3)
- Ημιφορτηγό (4)
- Φορτηγό (5)
- Λεωφορείο (6)

Διαχωρισμός κατευθύνσεων

Ο διαχωρισμός των δύο κατευθύνσεων έγινε με βάση το **διάνυσμα κατεύθυνσης** του κάθε οχήματος, την κάθε χρονική στιγμή. Σε περίπτωση όπου μεταξύ δύο διαδοχικών χρονικών στιγμών το όχημα παραμένει στάσιμο, εξετάστηκαν οι συντεταγμένες του οχήματος τις επόμενες χρονικές στιγμές μέχρι να βρεθεί η πρώτη επόμενη χρονική στιγμή κατά την οποία το όχημα έχει κινηθεί. Την τελευταία χρονική στιγμή της μελέτης, εφόσον ένα όχημα βρίσκεται εντός του εξεταζόμενου τμήματος, θεωρείται ότι η κατεύθυνση του παραμένει ίδια με την αμέσως προηγούμενη.

Στροφή αξόνων

Πραγματοποιήθηκε στροφή του καρτεσιανού συστήματος συντεταγμένων, έτσι ώστε οι **άξονες να συμφωνούν με αυτούς της οδού**. Δηλαδή ο άξονας y να είναι παράλληλος με τον άξονα της οδού και ο άξονας x να είναι κάθετος στην οδό. Κατά αυτόν τον τρόπο

γίνεται δυνατή η εύρεση των πλευρικών αποστάσεων και των αποστάσεων από το μπροστά και το πίσω όχημα του εξεταζόμενου οχήματος.

Υπολογισμός αποστάσεων

Η εύρεση αυτών των αποστάσεων είναι αναγκαία καθώς με βάση αυτές το μοντέλο θα κρίνει το επίπεδο ασφαλείας της θέσης του οχήματος σε σχέση με την γύρω κυκλοφορία. Συνεπώς θα πρέπει να αποτελούν **υπολογίσιμες από το ίδιο το όχημα αποστάσεις**, όπως και συμβαίνει με την χρήση από το όχημα αισθητήρων, χωρίς την ανάγκη δορυφορικής κάλυψης για την εύρεση συντεταγμένων.

Τέλος υπολογίσθηκαν οι **αποστάσεις μεταξύ του εξεταζόμενου οχήματος και του κοντινότερου οχήματος δεξιά, αριστερά, μπροστά και πίσω**, ή όποιος συνδυασμός αυτών υπήρχε την κάθε χρονική στιγμή γύρω από το όχημα. Ο συγκεκριμένος υπολογισμός έγινε για κάθε κατεύθυνση ξεχωριστά καθώς στη Λεωφόρο Αλεξάνδρας υπάρχει διαχωριστική νησίδα, οπότε τα δύο ρεύματα είναι πλήρως ανεξάρτητα. Για να είναι με βεβαιότητα ένα όχημα δεξιά, αριστερά, μπροστά ή πίσω του εξεταζόμενου αντίστοιχα, τέθηκαν κάποια συντηρητικά όρια απόκλισης από το κέντρο του οχήματος, καθώς τα διαθέσιμα δεδομένα αναπαριστούν το κάθε όχημα ως σημείο (αδιάστατο), οπότε δεν είναι γνωστές οι διαστάσεις του κάθε οχήματος. Αυτά τα όρια κρίθηκαν με βάση τον τύπο και τις τυπικές διαστάσεις των δύο υπό μελέτη οχημάτων. Για να βρίσκεται ένα όχημα δεξιά ή αριστερά του εξεταζόμενου ελέγχεται η απόκλιση των κέντρων τους κατά τον άξονα y , ενώ για να βρίσκεται μπροστά ή πίσω κατά τον άξονα x . Οι διαστάσεις αυτές παρουσιάζονται στον Πίνακα 2.

Πίνακας 2: Επιτρεπτή απόκλιση αποστάσεων

Τύποι Οχημάτων	Dx	Dy
Αυτοκίνητο - Μοτοσυκλέτα	0.8	3
Αυτοκίνητο - Αυτοκίνητο	1	4
Αυτοκίνητο - Ημιφορτηγό	1.2	4
Αυτοκίνητο - Φορτηγό	1.5	5
Αυτοκίνητο - Λεωφορείο	1.5	5

Τα **αυτοκίνητα και τα ταξί θεωρούνται ως ίδιος τύπος οχήματος** για την οριοθέτηση καθώς έχουν ίδιες διαστάσεις. Αναλόγως με το πρόσημο του Dx και του Dy και την κατεύθυνση που εξετάζεται κρίνεται αν κάθε όχημα είναι δεξιά ή αριστερά και μπροστά ή πίσω αντίστοιχα. Στην περίπτωση όπου βρεθούν περισσότερα του ενός οχήματος σε μία από τις πλευρές, λαμβάνεται υπόψιν το όχημα που απέχει την μικρότερη απόσταση από το εξεταζόμενο.

Αφότου ολοκληρωθεί η συγκεκριμένη διαδικασία για όλα τα δεδομένα, δημιουργείται ο **Master Table** στον οποίο περιέχονται όλες οι καταστάσεις κατά τις οποίες γύρω από το εκάστοτε εξεταζόμενο όχημα υπάρχει τουλάχιστον ένα άλλο όχημα. Έτσι αν το εξεταζόμενο όχημα είναι μόνο του στην περιοχή του δρόμου, δεν εξετάζεται καθώς ο

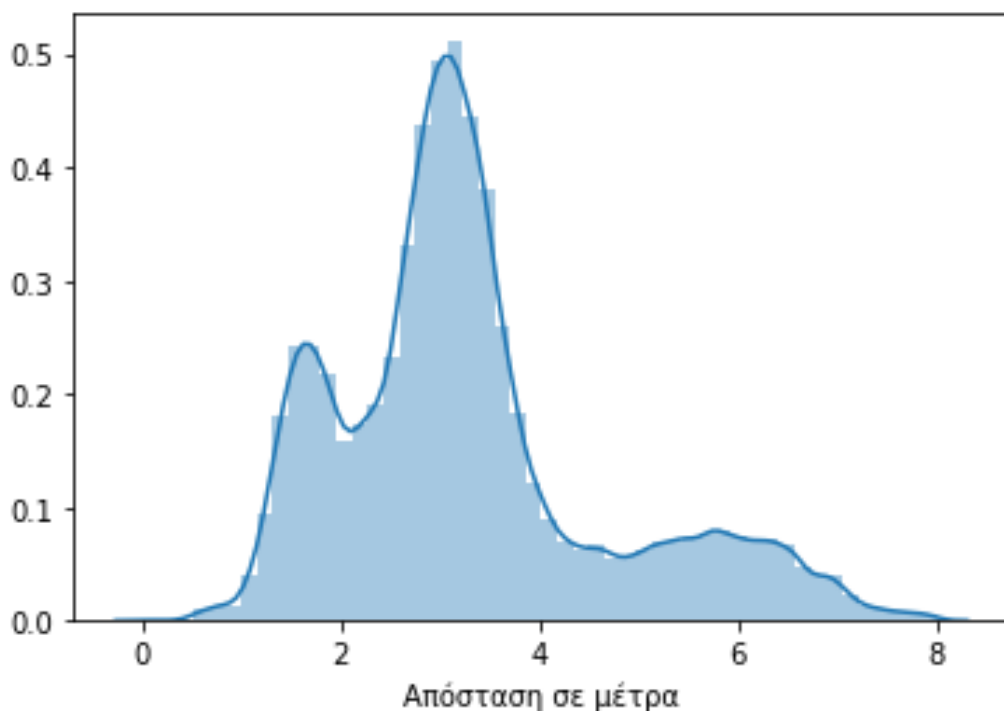
αλγόριθμός αφορά στην αξιολόγηση του επιπέδου ασφαλείας του εξεταζόμενου οχήματος σε σχέση με την υπόλοιπη κυκλοφορία. Κάθε γραμμή του τελικού πίνακα είναι ανεξάρτητη από τις υπόλοιπες και περιέχει πληροφορία για:

- τον τύπο του εξεταζόμενου και των γύρω οχημάτων
- τις ταχύτητες των οχημάτων και
- τις αποστάσεις

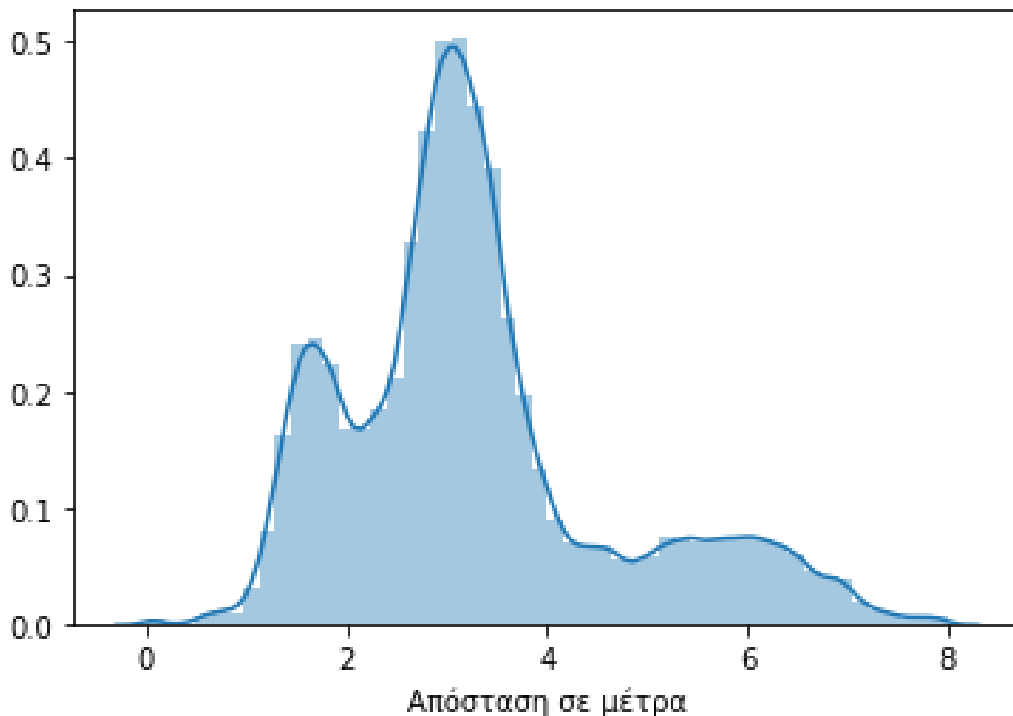
Οι ταχύτητες είναι αναγκαίες για τον υπολογισμό του χρόνου για σύγκρουση (**Time to Collision**, TTC) του εξεταζόμενου οχήματος με το μπροστινό όχημα και του πίσω οχήματος με το εξεταζόμενο. Μεγάλες διαφορές ταχυτήτων αποτελούν ένδειξη επιθετικής οδήγησης και πρέπει να ληφθούν υπόψιν στην αξιολόγηση του επιπέδου ασφαλείας και στην επιλογή της κατάλληλης ενέργειας.

4.4 ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Στα Διαγράμματα 3 έως 6 φαίνονται οι κατανομές **των αποστάσεων** του εξεταζόμενου με τα αριστερά, δεξιά, μπροστά και πίσω οχήματα αντίστοιχα. Μέσω της αξιολόγησης αυτών των διαγραμμάτων προκύπτουν ορισμένα συμπεράσματα.



Διάγραμμα 3: Διασπορά των αποστάσεων του εξεταζόμενου με το αριστερά όχημα

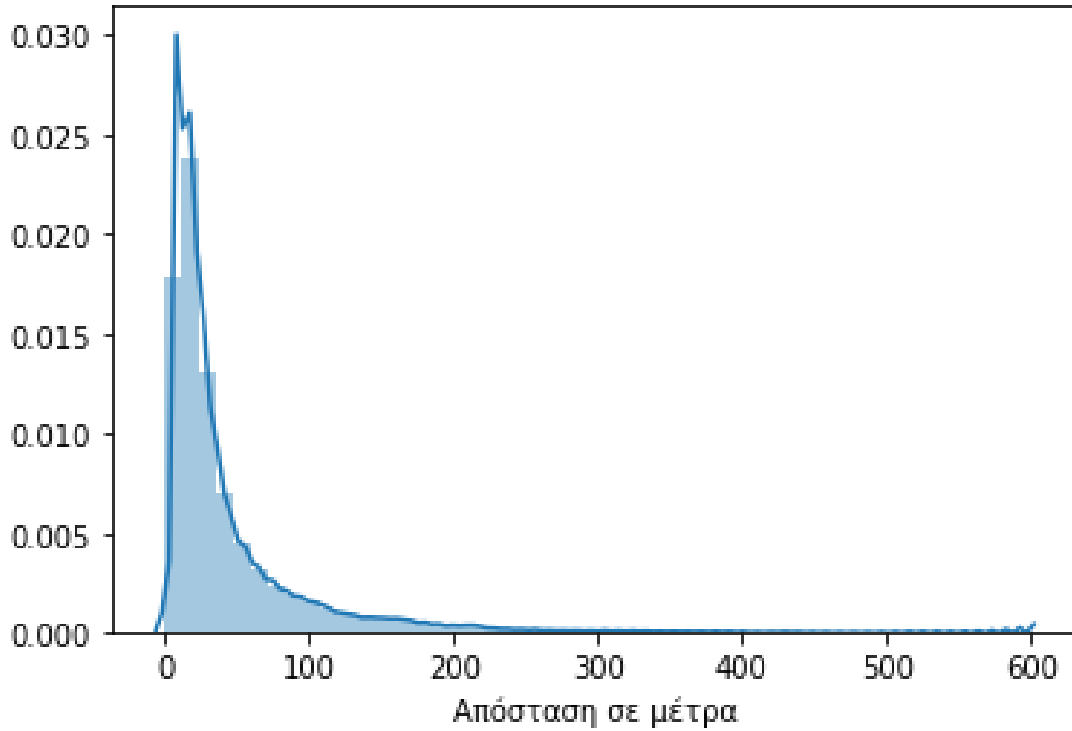


Διάγραμμα 4: Διασπορά των αποστάσεων του εξεταζόμενου με το δεξιό όχημα

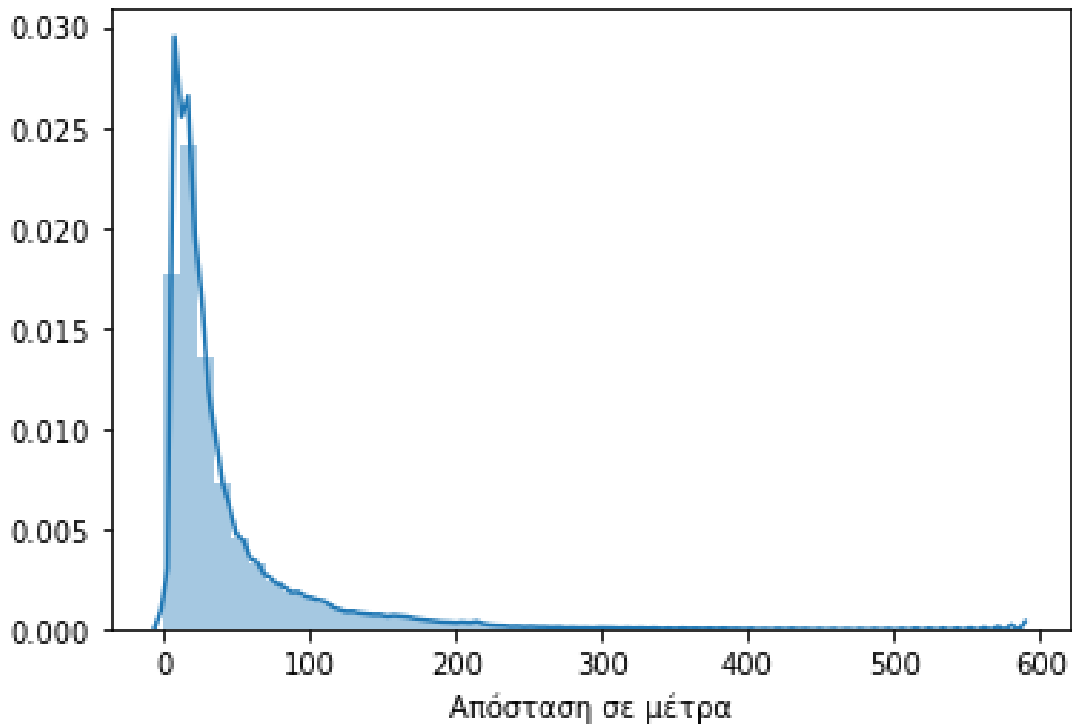
Παρατηρείται ότι το διάγραμμα των πλευρικών αποστάσεων είναι **παρόμοιο για τα δεξιά και τα αριστερά οχήματα** (Διάγραμμα 3 και Διάγραμμα 4), το οποίο μπορεί να οφείλεται σε διάφορους παράγοντες. Η πρώτη κορυφή του διαγράμματος η οποία εντοπίζεται γύρω στα 1.5 μέτρα οφείλεται στην ύπαρξη μεγάλου αριθμού μηχανοκίνητων δίκυκλων στην κυκλοφορία κατά την διάρκεια των μετρήσεων. Εφόσον οι μετρήσεις πραγματοποιήθηκαν σε αστικό περιβάλλον είναι αναμενόμενη η εμφάνιση μεγάλου αριθμού δίκυκλων.

Η δεύτερη κορυφή η οποία βρίσκεται περίπου στα 3 μέτρα **ανταποκρίνεται στον μεγάλο αριθμό αυτοκινήτων και ταξί**. Η απόσταση των τριών μέτρων ως μέση απόσταση μεταξύ αυτοκινήτων είναι αναμενόμενη λαμβάνοντας υπόψιν όχημα πλάτους 1.6-2 μέτρων και απόσταση 1 μέτρου κενή ανάμεσα στα οχήματα. Κατά την διάρκεια των μετρήσεων δεν παρατηρήθηκαν μεγάλοι κυκλοφοριακοί φόρτοι με φαινόμενα συμφόρησης, συνεπώς δεν χρειάστηκε οι οδηγοί να βρίσκονται σε πολύ μικρές αποστάσεις και τα δίκυκλα να κινούνται ανάμεσα σε οχήματα.

Όσον αφορά στα μπροστά και πίσω οχήματα (Διάγραμμα 5 και Διάγραμμα 6), η κορυφή του διαγράμματος κατανομής των αποστάσεων εμφανίζεται στα 20 μέτρα περίπου. Τα **οχήματα κινούνται σε ομάδες διατηρώντας μια ικανοποιητική απόσταση**, δεδομένων των κυκλοφοριακών συνθηκών. Για λόγους πληρότητας, διατηρήθηκαν όλες οι αποστάσεις, οι οποίες φτάνουν τα 600 μέτρα περίπου, όσο το μήκος του εξεταζόμενου οδικού τμήματος.



Διάγραμμα 5: Διασπορά των αποστάσεων του εξεταζόμενου με το μπροστά όχημα



Διάγραμμα 6: Διασπορά των αποστάσεων του εξεταζόμενου με το πίσω όχημα

5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 ΕΙΣΑΓΩΓΗ

Το κεφάλαιο αυτό περιλαμβάνει την αναλυτική περιγραφή της μεθοδολογίας που χρησιμοποιήθηκε και την παρουσίαση των αποτελεσμάτων της Διπλωματικής Εργασίας. Έπειτα από την συλλογή και την επεξεργασία των στοιχείων σε περιβάλλον *rython*, όπως περιεγράφηκε στο προηγούμενο κεφάλαιο, ακολουθεί **η εφαρμογή της μεθόδου της ενισχυτικής μάθησης**. Για την διαμόρφωση του προβλήματος με την μέθοδο **Q-Mάθησης**, πρέπει να προσδιοριστούν οι καταστάσεις του περιβάλλοντος της οδήγησης, οι διαθέσιμες ενέργειες του πράκτορα, η συνάρτηση ανταμοιβής για την προσαρμογή της πολιτικής και οι παράμετροι της ενισχυτικής μάθησης.

5.2 ΚΑΤΑΣΤΑΣΕΙΣ

Το **πρόβλημα της αναπαράστασης των καταστάσεων** στην ενισχυτική μάθηση είναι παρόμοιο με τα προβλήματα αντιπροσώπευσης χαρακτηριστικών και επιλογής χαρακτηριστικών στην μάθηση με επίβλεψη και χωρίς επίβλεψη. Στην εφαρμογή και σύγκριση των διαφόρων αλγορίθμων ενισχυτικής μάθησης τα περιβάλλοντα που χρησιμοποιούνται θα πρέπει να είναι αρκετά απλά, ώστε να είναι εύκολη η απεικόνιση όλων των καταστάσεων. Αυτό απλοποιεί τις εκτιμήσεις των αξιών σε βασικούς κυλιόμενους μέσους όρους σε έναν πίνακα, οι οποίοι είναι ευκολότερο να κατανοηθούν και να εφαρμοστούν.

Όστούσο τα **περισσότερα προβλήματα ελέγχου δεν μπορούν να αναπαρασταθούν με έναν συγκεκριμένο αριθμό καταστάσεων**. Αυτό οφείλεται κυρίως στην “*Κατάρα της Διαστασιμότητας*” (Curse of Dimensionality), όπου οι διαστάσεις ενός προβλήματος είναι πάρα πολλές και δύσκολα απεικονίζονται. Για παράδειγμα, στο πρόβλημα της οδήγησης υπάρχουν διάφορες αποστάσεις μεταξύ του οχήματος και της γύρω κυκλοφορίας, διάφορες ταχύτητες και θέσεις των οχημάτων και γωνίες κίνησης. Έτσι εφόσον τα μεγέθη είναι συνεχή, υπάρχουν άπειρες καταστάσεις οι οποίες περιγράφουν το περιβάλλον.

Επιπλέον, στην ενισχυτική μάθηση, πρέπει να ληφθεί υπόψιν και η **ιδιότητα Markov**. Θα πρέπει η κατάσταση να παρέχει αρκετές πληροφορίες, έτσι ώστε να μπορούν να προβλεφθούν με ακρίβεια οι επόμενες ανταμοιβές και οι επόμενες καταστάσεις δεδομένης μίας ενέργειας, χωρίς την ύπαρξη πληροφοριών για το παρελθόν. Οποιοσδήποτε παράγοντας που είναι ουσιαστικά τυχαίος μπορεί επίσης να αγνοηθεί.

Λαμβάνοντας υπόψιν όλα τα παραπάνω και για να απλοποιηθεί το πρόβλημα της Κατάρας της Διαστασιμότητας για τον προσδιορισμό των καταστάσεων λαμβάνονται υπόψιν οι εξής παράγοντες

- Ο χρόνος για σύγκρουση του εξεταζόμενου οχήματος με το μπροστινό όχημα
- Ο χρόνος για σύγκρουση του εξεταζόμενου οχήματος με το πίσω όχημα
- Οι αποστάσεις του εξεταζόμενου οχήματος με τα πλευρικά οχήματα

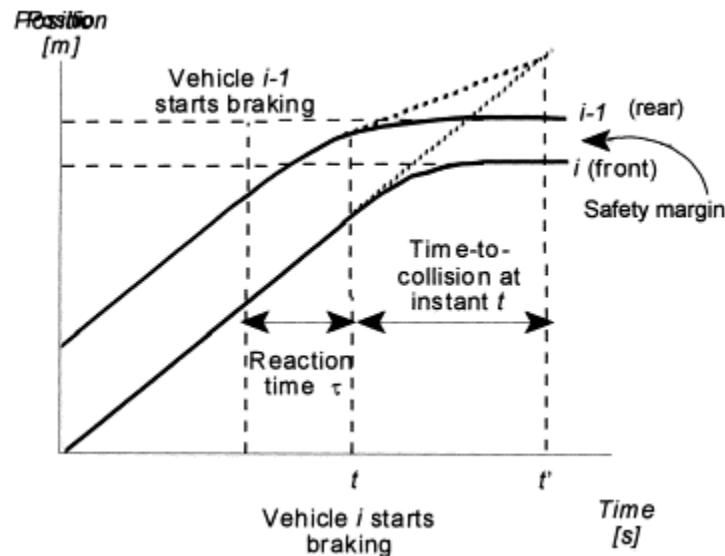
Όλα τα μεγέθη αυτά είναι εύκολα μετρήσιμα από αισθητήρες που τοποθετούνται στο εξεταζόμενο όχημα. Έτσι δεν απαιτείται επικοινωνία μεταξύ των οχημάτων (connected vehicles) και οι υπολογισμοί μπορούν να γίνουν σε πραγματικό χρόνο από το όχημα. **Το πρόγραμμα λοιπόν είναι εύκολα εφαρμόσιμο** από τα οχήματα αρκεί να υπάρχει ο απαραίτητος εξοπλισμός.

Χρόνος για Σύγκρουση

Ο όρος Χρόνος για Σύγκρουση (**Time to Collision**) αναφέρεται στον χρόνο που απαιτείται, συνήθως σε δευτερόλεπτα, μέχρις ότου δύο οχήματα διατηρώντας τις ταχύτητες τους συγκρουστούν. Ο Χρόνος για Σύγκρουση μεταξύ ενός οχήματος i την χρονική στιγμή t σε σχέση με ένα προπορευόμενο όχημα $i-1$ υπολογίζεται από τον τύπο:

$$TTC = \frac{X_{i-1}(t) - X_i(t) - l_i}{X_i(t) - X_{i-1}(t)} \quad (1)$$

Για κάθε $X_i(t) > X_{i-1}(t)$, όπου \dot{X}_i συμβολίζει την ταχύτητα, X την θέση και l το μήκος του οχήματος.



Εικόνα 11: Χρόνος για σύγκρουση μεταξύ δυο οχημάτων (Πηγή: Minderhoud et al. 2000)

Στην Εικόνα 11 αναπαρίσταται η ιδέα του Χρόνου για Σύγκρουση με δύο τροχιές οχημάτων. Εμφανίζεται μια κατάσταση όπου το όχημα $i-1$ που προπορεύεται φρενάρει. Μετά τον χρόνο αντίδρασης τ , ο οδηγός i ξεκινά μια ενέργεια ελέγχου. Η τιμή του Χρόνου για Σύγκρουση **TTC** μπορεί να υπολογιστεί μόνο όταν υπάρχει θετική διαφορά της ταχύτητας μεταξύ των οχημάτων και δίνεται την στιγμή τ που ο οδηγός αρχίζει να φρενάρει. Για τον υπολογισμό του μεγέθους θεωρείται ότι **η διαφορά των ταχυτήτων την χρονική στιγμή τ παραμένει σταθερή** κατά την διάρκεια των υποθετικών τροχιών

σύγκρουσης (όπως φαίνεται και στην Εικόνα 11). Όσο μεγαλύτερη είναι η τιμή τόσο πιο ασφαλής θεωρείται μια κατάσταση.

Η **κρίσιμη τιμή** του Χρόνου για Σύγκρουση έχει αποτελέσει θέμα διάφορων ερευνών με τα αποτελέσματα να μην συγκλίνουν. Οι Hist και Graham (1997) αναφέρουν ότι ένα μέτρο του Χρόνου για Σύγκρουση ίσο με τέσσερα δευτερόλεπτα θα μπορούσε να χρησιμοποιηθεί για την διάκριση μεταξύ περιπτώσεων όπου οι οδηγοί ακούσια βρίσκονται σε μια επικίνδυνη κατάσταση από τις περιπτώσεις όπου οι οδηγοί διατηρούν τον έλεγχο. Η μελέτη περιγράφει επίσης ένα εργαστηριακό πείραμα για τον σχεδιασμό ενός Συστήματος Αποφυγής Σύγκρουσης. Τα αποτελέσματα δείχνουν ότι ένα κριτήριο προειδοποίησης του Χρόνου για Σύγκρουση των τεσσάρων ή πέντε δευτερολέπτων είχε ως αποτέλεσμα πολλούς ψευδείς συναγερμούς. Η έρευνα αυτή δείχνει ότι τιμή του χρόνου ίση με τρία δευτερόλεπτα παρήγαγε τον ελάχιστο αριθμό λανθασμένων συναγερμών.

Οι Hogema και Janssen (1996) διερεύνησαν τη συμπεριφορά σε μία προσέγγιση μη υποστηριζόμενων και υποστηριζόμενων οδηγών σε ένα πείραμα προσομοιωτή οδήγησης. Από το πείραμα προέκυψε μια ελάχιστη τιμή του Χρόνου για Σύγκρουση ίση με 3.5 δευτερόλεπτα για τους οδηγούς χωρίς υποστήριξη και 2.4 δευτερόλεπτα για τους οδηγούς με υποστήριξη. **Η τιμή 2.4 θεωρείται ως κρίσιμη** σύμφωνα με αυτήν την έρευνα. Τέλος, ο VanDerHorst (1990) αναφέρει ακόμα μικρότερες κρίσιμες τιμές για τον Χρόνο για Σύγκρουση.

Με βάση τα παραπάνω για την εκπόνηση της Διπλωματικής Εργασίας θεωρείται ως κρίσιμη τιμή του Χρόνου για Σύγκρουση τα 2.4 δευτερόλεπτα. Η τιμή αυτή αποτελεί μια ενδιάμεση συντηρητική τιμή των διαφόρων ερευνών και είναι θεώρηση υπέρ της ασφαλείας. Συνεπώς, όταν εμφανίζεται **τιμή Χρόνου για Σύγκρουση μικρότερη από 2.4 δευτερόλεπτα, η κατάσταση θα θεωρείται επικίνδυνη.**

Χρόνος Αντίδρασης

Δεδομένου ότι το πρόγραμμα θα παρέχει καθοδήγηση για τον οδηγό μέσω μηνυμάτων, θα πρέπει να λαμβάνεται υπόψιν και ο **χρόνος αντίδρασης του οδηγού**. Αυτό διότι ο οδηγός από την στιγμή που το σύστημα θα μεταδώσει το μήνυμα χρειάζεται ένα χρονικό διάστημα μέχρι να προβεί σε ενέργειες. Έτσι θα πρέπει να λαμβάνεται υπόψιν ο χρόνος αντίδρασης του οδηγού για να αποφεύγονται οι κρίσιμες καταστάσεις.

Οι περισσότερες έρευνες εκτιμούν τον χρόνο αντίδρασης με βάση πειράματα σε εσωτερικό χώρο με χρήση προσομοιωτή. Οι Johansson και Rumar (1971), διερεύνησαν τον χρόνο για αντίδραση φρεναρίσματος και κατέληξαν στο ότι οι τιμές του χρόνου αντίδρασης **κυμαίνονται από 0.4 μέχρι 2.7** δευτερόλεπτα με μέση τιμή και τυπική απόκλιση 1.01 και 0.37 δευτερόλεπτα, αντίστοιχα. Ωστόσο δεδομένου ότι οι οδηγοί γνώριζαν ότι συμμετείχαν σε ένα πείραμα για τον υπολογισμό του χρόνου αντίδρασης και χρησιμοποιούνταν ένα ήχος για διέγερση, οι τιμές μπορεί να μην είναι αντικειμενικές.

Μια πρόσφατη έρευνα που χρησιμοποιεί συνδυασμό οδήγησης σε προσομοιωτή και σε πραγματικό περιβάλλον, έδειξε ότι ο μέσος χρόνος αντίδρασης των οδηγών σε ένα αναμενόμενο συμβάν σε πραγματικό περιβάλλον είναι 0.42 δευτερόλεπτα και η τυπική απόκλιση 0.14 δευτερόλεπτα (Magister et al., 2005). Ταυτόχρονα, η μέση τιμή του χρόνου

αντίδρασης σε έναν **μη αναμενόμενο κίνδυνο είναι 1.1 δευτερόλεπτο** και σε προσομοιωτή 0.9. Σε πραγματική κυκλοφορία οι χρόνοι αντίδρασης οδηγού σε προσδοκόμενα και απροσδόκητα ερεθίσματα διαφέρουν.

Ο Lerner (1995) κατά τη διερεύνηση του χρόνου αντίδρασης δεν ενημέρωσε τους οδηγούς ότι συμμετέχουν σε έρευνα του χρόνου αντίδρασης. Όταν λοιπόν ένας οδηγός άγγιζε την ταχύτητα των 64 χλμ/ώρα, ένα μεγάλο βαρέλι σύγκρουσης απελευθερωνόταν σε απόσταση 61 μέτρων μπροστά από το όχημα. Οι τιμές του χρόνου αντίδρασης κυμαίνονται μεταξύ 0.7 και 2.5 δευτερόλεπτα με μεσαία τιμή, μέση τιμή και τυπική απόκλιση 1.44, 1.51 και 0.39 δευτερόλεπτα αντίστοιχα.

Ο **χρόνος αντίδρασης λαμβάνεται ίσος με 1.5 δευτερόλεπτα** καθώς αποτελεί μια ενδιάμεση τιμή, αντιπροσωπευτική των περισσότερων οδηγών. Αυτός θα πρέπει να προστεθεί στον κρίσιμο Χρόνο για Σύγκρουση, έτσι ώστε ο οδηγός να λαμβάνει το μήνυμα και να προλαβαίνει να αντιδράσει πριν ο Χρόνος για Σύγκρουση γίνει κρίσιμος.

Τελικά ο χρόνος για σύγκρουση που θα λαμβάνεται υπόψιν για την ειδοποίηση του οδηγού τυχόν επικίνδυνης κατάστασης θα είναι ίσος με **3.9 δευτερόλεπτα**.

Κρίσιμες Πλευρικές Αποστάσεις

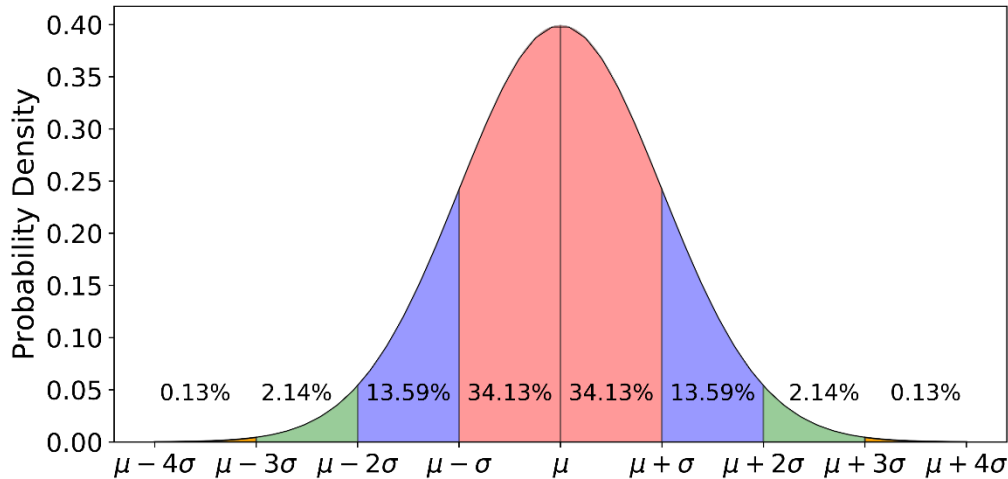
Ως κρίσιμη πλευρική απόσταση ορίζεται η απόσταση μεταξύ δύο οχημάτων τα οποία βρίσκονται το ένα δίπλα στο άλλο, πέραν την οποίας υπάρχει μεγάλος κίνδυνος πλευρικής σύγκρουσης των οχημάτων. Οι κρίσιμες πλευρικές αποστάσεις υπολογίζονται για τους συνδυασμούς των αυτοκινήτων με τα άλλα οχήματα. Οι 5 κατηγορίες οχημάτων είναι

1. Δίκυκλο
2. Αυτοκίνητο
3. Ημιφορτηγό
4. Φορτηγό
5. Λεωφορείο

Οι κρίσιμες πλευρικές αποστάσεις προκύπτουν για τους συνδυασμούς των οχημάτων με βάση την κατανομή όλων των πλευρικών αποστάσεων των συνδυασμών που παρατηρήθηκαν στα δεδομένα. **Οι κατανομές** προσεγγίζουν την κανονική κατανομή οπότε θεωρείται για τον υπολογισμό των κρίσιμων πλευρικών αποστάσεων ότι **ακολουθούν κανονική τιμή** (Εικόνα 12).

Οι πλευρικές αποστάσεις δεξιά και αριστερά των οχημάτων είναι σχεδόν ίδιες οπότε οι κρίσιμες αποστάσεις μεταξύ του εξεταζόμενου οχήματος και των **δεξιά και αριστερά οχημάτων θεωρούνται ίδιες**. Αυτό συμβαίνει διότι όταν ένα όχημα προσεγγίζει δύο άλλα οχήματα και εισέρχεται στην μεταξύ των δύο άλλων οχημάτων κενή λωρίδα, ο οδηγός προσπαθεί να εισέλθει στο μέσο της απόστασης των δύο άλλων οχημάτων. Επίσης όταν τρία οχήματα κινούνται σε διπλανές λωρίδες, ο οδηγός του μεσαίου οχήματος έχει καλύτερη εποπτεία για την απόσταση του με το δεξιά όχημα, ενώ ο αριστερός οδηγός έχει για την απόσταση μεταξύ του αριστερού οχήματος και του μεσαίου. Εν τέλει, οι πλευρικές αποστάσεις καταλήγουν να είναι παρόμοιες διότι η αριστερή απόσταση του μεσαίου οχήματος αποτελεί την δεξιά απόσταση του αριστερού οχήματος, και η δεξιά απόσταση του μεσαίου οχήματος αποτελεί την αριστερή απόσταση του δεξιά οχήματος.

Οι τελικές κρίσιμες αποστάσεις προκύπτουν από την **μέση τιμή της κατανομής αφαιρώντας την τυπική απόκλιση**. Αυτές οι τιμές ανταποκρίνονται στο 84.14% της κατανομής, δηλαδή το 15.86% των οχημάτων διατηρούν κρίσιμη πλευρική απόσταση. Το ποσοστό αυτό επιλέχθηκε συντηρητικά για να ληφθεί υπόψιν η διακύμανση του πλάτους των οχημάτων, καθώς οι διαστάσεις των οχημάτων διαφέρουν και οι διαθέσιμες μετρήσεις είναι από κέντρο σε κέντρο οχήματος και όχι καθαρές πλευρικές αποστάσεις.



Εικόνα 12: Κανονική Κατανομή (Πηγή: Galarnyk 2018)

Με βάση αυτό το ποσοστό **οι κρίσιμες πλευρικές αποστάσεις** είναι

1. Αυτοκίνητο – Μηχανοκίνητο Δίκυκλο: 1.350 m
2. Αυτοκίνητο – Αυτοκίνητο: 1.925 m
3. Αυτοκίνητο – Ημιφορτηγό: 2.382 m
4. Αυτοκίνητο – Φορτηγό: 2.556 m
5. Αυτοκίνητο – Λεωφορείο: 2.466 m

Σύμφωνα με αυτές ο αλγόριθμος θα κρίνει τις πλευρικές αποστάσεις μεταξύ του εξεταζόμενου οχήματος και των πλευρικών οχημάτων. Οι τιμές που προκύπτουν ανταποκρίνονται στις μέσες διαστάσεις των οχημάτων. Η πλευρική απόσταση με μηχανοκίνητο δίκυκλο είναι λίγο μεγάλη, αλλά αυτό οφείλεται στην έλλειψη έντονης κυκλοφοριακής συμφόρησης κατά τις μετρήσεις οπότε υπήρχε επαρκής χώρος για την κίνηση των δίκυκλων και δεν εμφανιζόταν κίνηση δίκυκλων μεταξύ των αυτοκινήτων. Επίσης, η πλευρική απόσταση μεταξύ αυτοκινήτων είναι σχετικά μικρή το οποίο μπορεί να οφείλεται στο ότι τα αυτοκίνητα εντός πόλης είναι μικρών διαστάσεων με μικρά πλάτη. Ομοίως η απόσταση μεταξύ αυτοκινήτου και ημιφορτηγού είναι μικρή λόγω του τρόπου με τον οποίο έγινε ο διαχωρισμός των τύπων των οχημάτων κατά την επεξεργασία των στοιχείων. Τέλος, οι αποστάσεις μεταξύ αυτοκινήτων και φορτηγών και λεωφορείων είναι αναμενόμενες.

Ορισμός Καταστάσεων

Ο ορισμός των καταστάσεων γίνεται σε **δύο στάδια**. Στο πρώτο στάδιο γίνεται διαχωρισμός με βάση την κυκλοφορία μπροστά και πίσω από το εξεταζόμενο όχημα, δηλαδή με βάση τα οχήματα που προπορεύονται του εξεταζόμενου και που το ακολουθούν. Στο δεύτερο στάδιο γίνεται διαχωρισμός με βάση την κυκλοφορία στα πλευρικά του οχήματος.

Στο Διάγραμμα 7 παρουσιάζεται το πρώτο στάδιο καθορισμού της κατάστασης στην οποία βρίσκεται το περιβάλλον. Αν λοιπόν μπροστά από το εξεταζόμενο όχημα υπάρχει κάποιο όχημα, τότε διερευνάται ο Χρόνος για Σύγκρουση του εξεταζόμενου οχήματος με το μπροστινό όχημα (Time to Collision-Front Vehicle, **TTCF**), ο οποίος θεωρείται κρίσιμος όταν η τιμή του είναι μικρότερη από 3.9 δευτερόλεπτα. Στην συνέχεια, αν υπάρχει πίσω όχημα τότε εξετάζεται ο Χρόνος για Σύγκρουση του πίσω οχήματος με το εξεταζόμενο όχημα (Time to Collision-Back Vehicle, **TTCB**) που επίσης θεωρείται κρίσιμος για τιμές μικρότερες των 3.9 δευτερολέπτων. Αυτό συμβαίνει διότι σε ορισμένες περιπτώσεις, το εξεταζόμενο όχημα ίσως χρειαστεί να επιβραδύνει σε περίπτωση που βρίσκεται σε μικρή απόσταση από το μπροστινό όχημα, σε σχέση με τις ταχύτητες των δύο οχημάτων. Αντίθετα, σε περίπτωση που το εξεταζόμενο όχημα κινδυνεύει να συγκρουστεί με το πίσω όχημα, τότε θα πρέπει να επιταχύνει για να το αποφύγει. Παρόλο που η επιτάχυνση είναι αντίθετη στο αίσθημα ασφάλειας του ανθρώπου, ορισμένες φορές αποτελεί την απαραίτητη ενέργεια για την αποφυγή ατυχήματος. Με βάση τα παραπάνω κριτήρια προκύπτουν **εννέα επίπεδα διαχωρισμού** των καταστάσεων του περιβάλλοντος. Στο δεύτερο στάδιο ο διαχωρισμός γίνεται με βάση την πλευρική κυκλοφορία

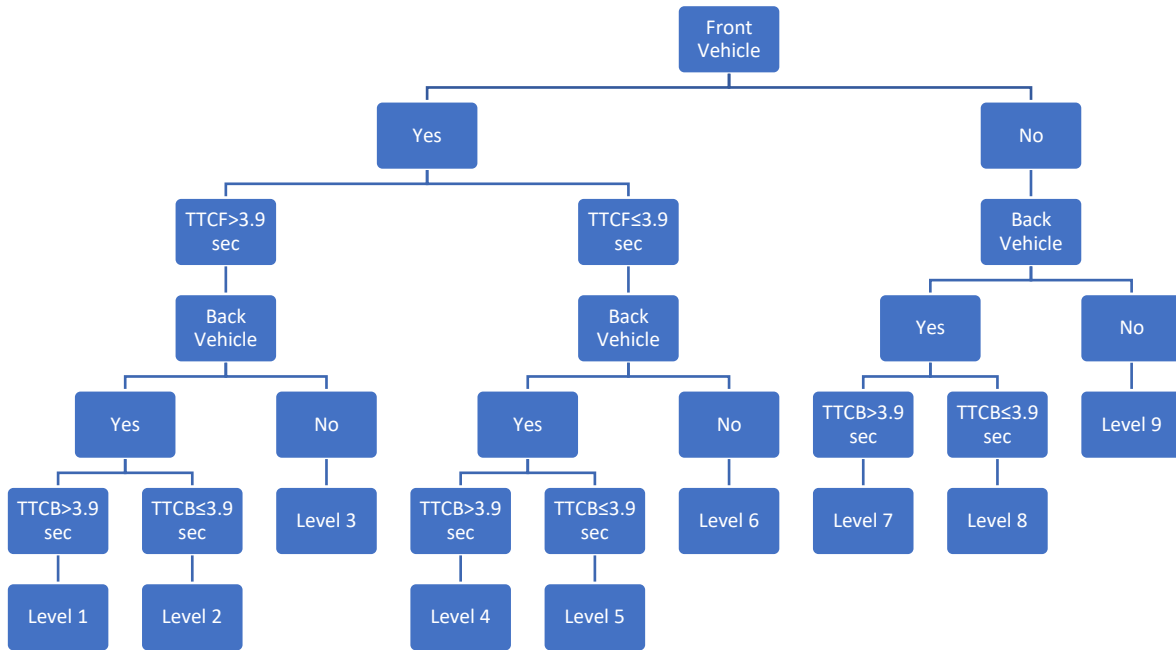
Στο Διάγραμμα 8 παρουσιάζεται το δεύτερο στάδιο διαχωρισμού των καταστάσεων. Εξετάζεται η ύπαρξη οχημάτων δεξιά και αριστερά του εξεταζόμενου οχήματος. Οι υποπεριπτώσεις είναι

1. Κανένα πλευρικό όχημα
2. Ένα πλευρικό όχημα
3. Δυο πλευρικά οχήματα

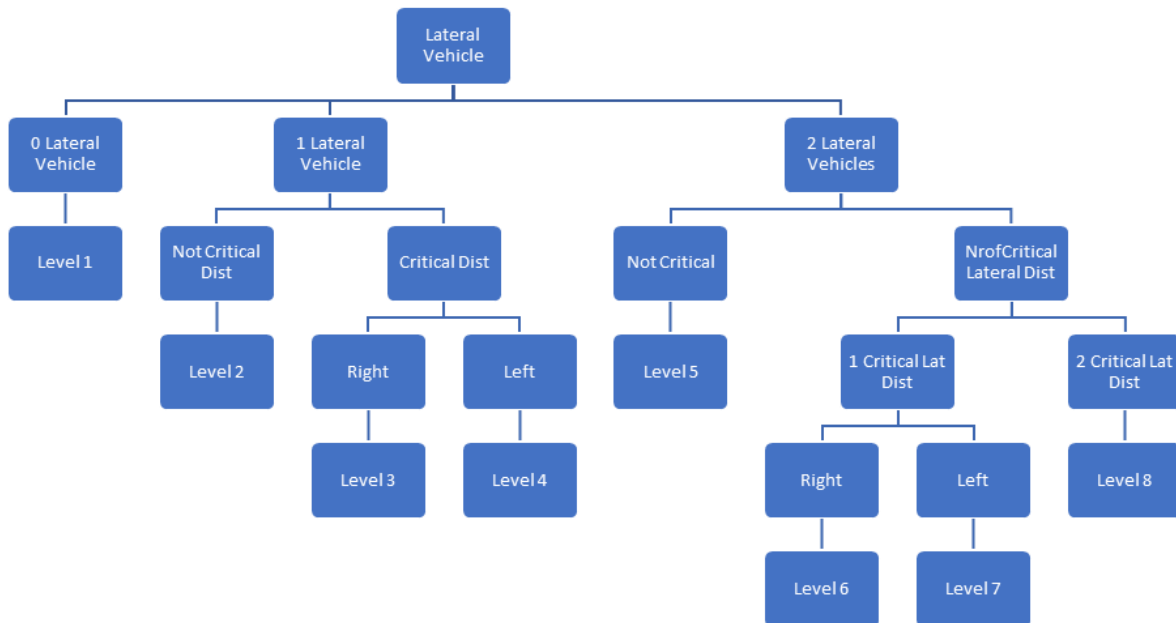
Αν σε κάποια πλευρά του εξεταζόμενου οχήματος υπάρχουν περισσότερα του ενός οχήματα τότε εξετάζεται το όχημα με την μικρότερη πλευρική απόσταση από το εξεταζόμενο. Στην περίπτωση όπου υπάρχουν πλευρικά οχήματα, διερευνάτε η κρισιμότητα των αποστάσεων αυτών των οχημάτων από το εξεταζόμενο. Έτσι προκύπτουν έξι επίπεδα διαχωρισμού. Για την περίπτωση όπου υπάρχει πλευρικό όχημα σε κρίσιμη απόσταση στη μια από τις δύο πλευρές του εξεταζόμενου οχήματος θα πρέπει να γίνει σαφές σε ποια πλευρά βρίσκεται το κρίσιμο όχημα γι' αυτό εν τέλει προκύπτουν **οκτώ επίπεδα διαχωρισμού**. Αυτό είναι αναγκαίο καθώς στην περίπτωση όπου υπάρχουν κρίσιμα οχήματα και στις δύο πλευρές ή σε καμία από τις δύο πλευρές δεν υπάρχει ανάγκη διαχωρισμού δεξιά και αριστερά. Αντίθετα, όταν υπάρχει κρίσιμο όχημα σε μια πλευρά είναι αναγκαίο για την καλύτερη λειτουργία του αλγόριθμου να διαχωρίζονται οι περιπτώσεις σε δεξιά και αριστερά ανάλογα με το που βρίσκεται το κρίσιμο όχημα.

Συνολικά το περιβάλλον μπορεί να βρεθεί σε **72 καταστάσεις** καθώς κάθε ένα από τα εννέα επίπεδα του Διαγράμματος 7 διαχωρίζεται σε οκτώ επιπλέον επίπεδα με βάση το Διάγραμμα 8.

5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ



Διάγραμμα 7: Έλεγχος μπροστά-πίσω οχημάτων



Διάγραμμα 8: Έλεγχος πλευρικής κυκλοφορίας

5.3 ΕΝΕΡΓΕΙΕΣ

Για την υλοποίηση του αλγορίθμου απαιτείται ο προσδιορισμός των διαθέσιμων ενεργειών που μπορεί να επιλέξει ο πράκτορας. Αυτές προκύπτουν μέσα από την διαδικασία της οδήγησης και διακρίνονται σε επέμβαση στην **διεύθυνση** οδήγησης και επέμβαση στην **ταχύτητα** του οχήματος. Έτσι ως προς την ταχύτητα υπάρχει η δυνατότητα επιτάχυνσης και επιβράδυνσης και ως προς την διεύθυνση υπάρχει η δυνατότητα κίνησης προς τα αριστερά ή προς τα δεξιά σε σχέση με την τωρινή διεύθυνση του οχήματος. Σε κάθε περίπτωση υπάρχει και η δυνατότητα να μην γίνει καμία επέμβαση σε αυτά τα χαρακτηριστικά, δηλαδή το όχημα να συνεχίσει με την ίδια ταχύτητα και προς την ίδια διεύθυνση. Με βάση τα παραπάνω προκύπτουν οι εξής ενέργειες:

1. Τίποτα-Τίποτα
2. Τίποτα-Επιτάχυνση
3. Τίποτα-Επιβράδυνση
4. Αριστερά-Τίποτα
5. Δεξιά-Τίποτα
6. Αριστερά-Επιτάχυνση
7. Αριστερά-Επιβράδυνση
8. Δεξιά-Επιτάχυνση
9. Δεξιά-Επιβράδυνση

Όπου ο πρώτος όρος κάθε ενέργειας αναφέρεται στην διεύθυνση και ο δεύτερος στην ταχύτητα.

Σε κάθε χρονικό βήμα ο πράκτορας έχει την δυνατότητα να επιλέξει μέσα από αυτές τις ενέργειες, ανάλογα με το ποια υπολογίζεται ότι θα επιφέρει την μεγαλύτερη επιστροφή. Αφού επιλέξει μια ενέργεια λαμβάνει μια ανταμοιβή και υπολογίζονται εκ νέου οι αξίες της κάθε ενέργειας.

Στόχος του πράκτορα είναι μακροπρόθεσμα να καταλήξει στις βέλτιστες ενέργειες για κάθε κατάσταση και κατ' επέκταση στην μεγιστοποίηση της αθροιστικής ανταμοιβής.

5.4 ΣΥΝΑΡΤΗΣΗ ΑΝΤΑΜΟΙΒΗΣ

Αφότου ο πράκτορας επιλέξει ενέργεια, η συνάρτηση ανταμοιβής αξιολογεί αυτή την επιλογή και επιστρέφει μια ανταμοιβή. Η αξιολόγηση αυτή γίνεται με τον εξής τρόπο

- Αν ο πράκτορας επιλέξει **την κατάλληλη ενέργεια** για να οδηγήσει το όχημα σε μια ασφαλέστερη κατάσταση ανταμείβεται με **μία μονάδα**. Μισή μονάδα για την σωστή επιλογή ως προς τον ένα άξονα κίνησης (μπροστά – πίσω) και μισή για την σωστή επιλογή ως προς τον άλλο άξονα (δεξιά – αριστερά)
- Αν ο πράκτορας επιλέξει **την σωστή ενέργεια ως προς τον ένα άξονα κίνησης του οχήματος και λάθος ενέργεια ως προς τον άλλο** ανταμείβεται με **μείον μισή μονάδα** καθώς με αυτόν τον τρόπο καθοδηγείται προς την σωστή πολιτική χωρίς να λαμβάνει την μέγιστη δυνατή ανταμοιβή. Έτσι λαμβάνει μισή μονάδα για την σωστή επιλογή ενέργειας ως προς τον ένα άξονα και μείον μια μονάδα για την λάθος επιλογή ως προς τον άλλο άξονα. Η συνολική ανταμοιβή είναι αρνητική

καθώς η ενέργεια που επέλεξε ο πράκτορας εξακολουθεί να είναι λάθος και δεν βοηθάει στην ασφάλεια της οδήγησης

- Αν ο πράκτορας επιλέξει **λάθος ενέργειες και ως προς τους δύο άξονες** κίνησης ανταμείβεται με **μείον δύο μονάδες**. Αυτό διότι η ενέργεια που επέλεξε οδηγεί το όχημα σε μια πιθανώς πιο επικίνδυνη κατάσταση από αυτή που ήταν προηγουμένως.

Η σωστή ενέργεια σε κάθε κατάσταση προκύπτει με βάση την λογική. Έτσι όταν υπάρχει κρίσιμος χρόνος για σύγκρουση μεταξύ του εξεταζόμενου και του μπροστινού οχήματος ο οδηγός θα πρέπει να μειώσει ταχύτητα, ενώ, όταν ο χρόνος για σύγκρουση είναι κρίσιμος μεταξύ του εξεταζόμενου και του πίσω οχήματος, ο οδηγός θα πρέπει να αυξήσει ταχύτητα, το οποίο παρόλο που πάει ενάντια στο αίσθημα της ασφάλειας του ανθρώπου αποτελεί την κατάλληλη ενέργεια υπέρ της ασφαλείας. Όταν υπάρχει όχημα σε κρίσιμη απόσταση δεξιά του εξεταζόμενου τότε το εξεταζόμενο όχημα θα πρέπει να κινηθεί προς τα αριστερά, ενώ όταν υπάρχει στα αριστερά θα πρέπει να κινηθεί προς τα δεξιά.

Στις ακραίες περιπτώσεις ο οδηγός θα πρέπει να παραμένει όσο το δυνατόν ανενεργός. Έτσι όταν υπάρχει όχημα σε κρίσιμη απόσταση και από τα αριστερά και από τα δεξιά του εξεταζόμενου τότε ο οδηγός δεν θα πρέπει να μετακινηθεί ούτε δεξιά ούτε αριστερά καθώς κινδυνεύει να συγκρουστεί. Ομοίως όταν ο χρόνος για σύγκρουση είναι κρίσιμος τόσο μεταξύ του εξεταζόμενου με το μπροστά όχημα όσο και το πίσω τότε επίσης ο οδηγός θα πρέπει να προσπαθεί να πετύχει την ενδιάμεση ταχύτητα του μπροστά και του πίσω οχήματος και να μην επιταχύνει ούτε να επιβραδύνει έντονα.

Αρχικές συνθήκες

Δεδομένου ότι η Q-Μάθηση είναι ένας επαναληπτικός αλγόριθμος, προϋποθέτει εμμέσως μια αρχική συνθήκη πριν από την πρώτη ενημέρωση. Οι υψηλές αρχικές τιμές, γνωστές και ως “αισιόδοξες αρχικές συνθήκες”, μπορούν να ενθαρρύνουν την εξερεύνηση καθώς ανεξάρτητα από την επιλογή της ενέργειας, ο κανόνας ενημέρωσης θα προκαλέσει χαμηλότερες αξίες από την άλλη εναλλακτική, αυξάνοντας έτσι την πιθανότητα επιλογής τους. Η πρώτη ανταμοιβή μπορεί να χρησιμοποιηθεί για την επαναφορά των αρχικών συνθηκών. Συμφωνά με αυτή την λογική, την πρώτη φορά που θα επιλεγεί μια ενέργεια, η ανταμοιβή χρησιμοποιείται για να ορίσει την τιμή του Q. Ένα μοντέλο που ενσωματώνει την επαναφορά των αρχικών συνθηκών (RIC) αναμένεται να προβλέψει τη συμπεριφορά των συμμετεχόντων καλύτερα από ένα μοντέλο που υποθέτει οποιαδήποτε αυθαίρετη αρχική συνθήκη (AIC). Το RIC φαίνεται να είναι συνεπές με την ανθρώπινη συμπεριφορά σε επαναλαμβανόμενα πειράματα δυαδικής επιλογής. Οι αρχικές συνθήκες για την εφαρμογή της μεθόδου στην συγκεκριμένη διπλωματική εργασία θεωρούνται μηδενικές. Δηλαδή ο Πίνακας-Q αποτελείται από μηδενικά για κάθε συνδυασμό κατάστασης-ενέργειας.

5.5 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Για την εφαρμογή και εκπαίδευση της μεθόδου χρησιμοποιήθηκε η βάση δεδομένων όπως αυτή διαμορφώθηκε στο κεφάλαιο της Επεξεργασίας Στοιχείων. Ο πίνακας των στοιχείων αποτελείται από **274976 γραμμές και 14 στήλες**. Η κάθε γραμμή είναι ανεξάρτητη και αποτελεί ένα χρονικό βήμα εκπαίδευσης. Από το σύνολο των δεδομένων το 80% χρησιμοποιήθηκε για εκπαίδευση και το 20% για έλεγχο της ποιότητας των αποτελεσμάτων. Δηλαδή ο αλγόριθμος **εκπαιδεύεται χρησιμοποιώντας 220000** δείγματα και **ελέγχεται χρησιμοποιώντας 54976**. Στις 14 στήλες περιλαμβάνονται όλες οι απαραίτητες πληροφορίες για την αξιολόγηση της κατάστασης από τον αλγόριθμο και είναι της μορφής

Στήλη 1	Στήλη 2	Στήλη 3	Στήλη 4	Στήλη 5	Στήλη 6	Στήλη 7
Τύπος Εξεταζόμενου Οχήματος	Ταχύτητα Εξεταζόμενου Οχήματος	Τύπος Δεξιά Οχήματος	Ταχύτητα Δεξιά Οχήματος	Απόσταση Δεξιά Οχήματος από Εξεταζόμενο	Τύπος Αριστερά Οχήματος	Ταχύτητα Αριστερά Οχήματος

Στήλη 8	Στήλη 9	Στήλη 10	Στήλη 11	Στήλη 12	Στήλη 13	Στήλη 14
Απόσταση Αριστερά Οχήματος από Εξεταζόμενο	Τύπος Μπροστά Οχήματος	Ταχύτητα Μπροστά Οχήματος	Απόσταση Μπροστά Οχήματος από Εξεταζόμενο	Τύπος Πίσω Οχήματος	Ταχύτητα Πίσω Οχήματος	Απόσταση Πίσω Οχήματος από Εξεταζόμενο

Αυτά αποτελούν στοιχεία τα οποία ευκολά υπολογίζονται από τους αισθητήρες του οχήματος έτσι η μέθοδος μπορεί να εφαρμοστεί σε οχήματα χωρίς ιδιαίτερες τροποποιήσεις.

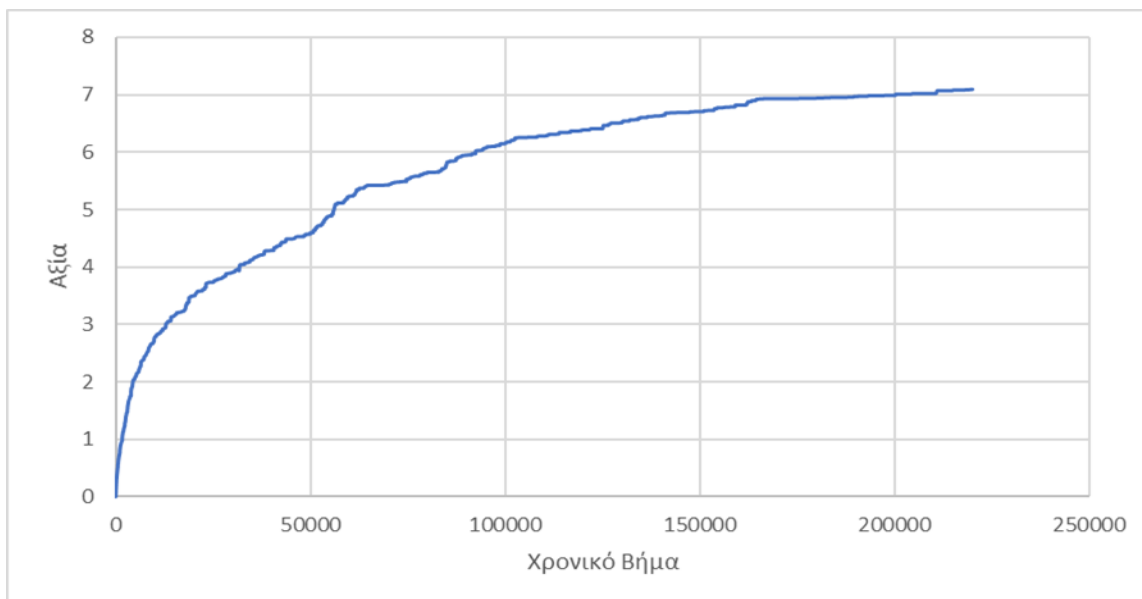
5.6 ΑΠΟΤΕΛΕΣΜΑΤΑ

Για την εξαγωγή του βέλτιστου αλγορίθμου πραγματοποιήθηκαν δοκιμές για διάφορους συνδυασμούς παραμέτρων. Οι τρεις παράμετροι της Q-Μάθησης είναι ο ρυθμός εκμάθησης lr , ο εκπτώτικος παράγοντας γ και ο βαθμός εξερεύνησης e . Το εύρος τιμών αυτών των παραμέτρων είναι από μηδέν μέχρι ένα. Τα 6 τελικά σενάρια τα οποία θα αξιολογηθούν είναι:

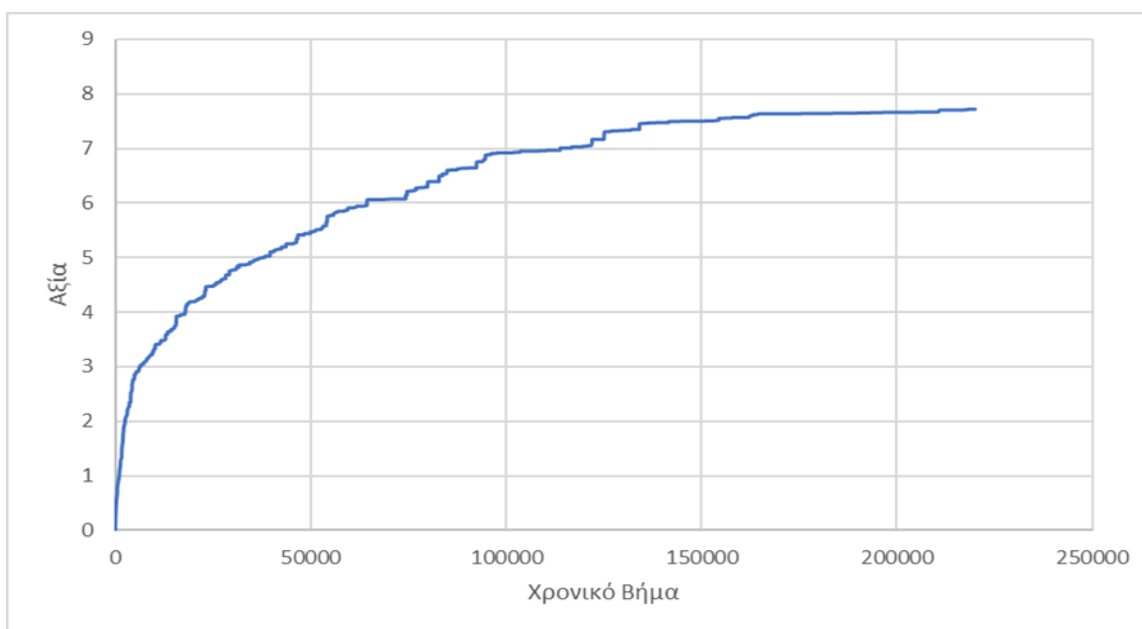
1. $\gamma=0.9$ $lr=0.2$ $e=0.4$
2. $\gamma=0.9$ $lr=0.5$ $e=0.4$
3. $\gamma=0.9$ $lr=0.2$ $e=0.2$
4. $\gamma=0.9$ $lr=0.2$ $e=0.5$
5. $\gamma=0.9$ $lr=0.2$ $e=0.7$
6. $\gamma=0.7$ $lr=0.2$ $e=0.2$

Ανάλυση Ρυθμού Εκμάθησης

Ο ρυθμός εκμάθησης καθορίζει σε τι βαθμό η καινούργια γνώση θα υπερισχύσει της παλιάς. Μεγαλύτερος ρυθμός εκμάθησης οδηγεί τον αλγόριθμο σε σύγκλιση γρηγορότερα από ένα μικρότερο ρυθμό. Δοκιμές πραγματοποιήθηκαν για διάφορες τιμές του ρυθμού εκμάθησης. Με βάση τα σενάρια 1 και 2 όπου διατηρούνται σταθερές οι άλλες δύο παράμετροι και αλλάζει ο ρυθμός εκμάθησης προκύπτουν τα Διαγράμματα 9 και 10 αντίστοιχα που παρουσιάζουν την διαχρονική μεταβολή της μέσης τιμής των βέλτιστων ενεργειών όλων των καταστάσεων.



Διάγραμμα 9: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $lr=0.2$

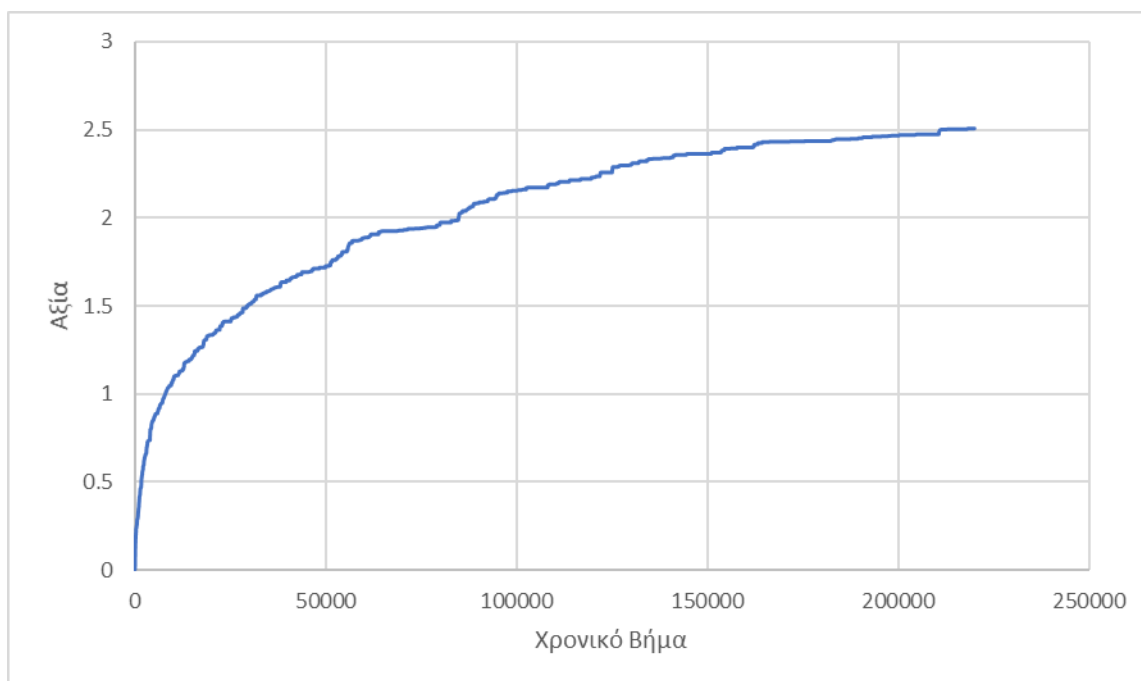


Διάγραμμα 10: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $lr=0.5$

Η μεγαλύτερη τιμή του ρυθμού εκμάθησης οδηγεί τον αλγόριθμο στο να μάθει πιο γρήγορα όπως γίνεται εμφανές και από την μεγαλύτερη κλίση που έχει ο αρχικός κλάδος του διαγράμματος. Ωστόσο πολλές φορές μεγάλη τιμή της παραμέτρου μπορεί να οδηγήσει σε μοντέλο που δύσκολα συγκλίνει, σε αντίθεση με μικρές τιμές όπου το μοντέλο με πιο αργό ρυθμό θα φτάσει σε σύγκλιση. Η διαφορά της αξίας στην οποία συγκλίνει ο κάθε συνδυασμός δεν αποτελεί ένδειξη της ποιότητας των τελικών αποτελεσμάτων.

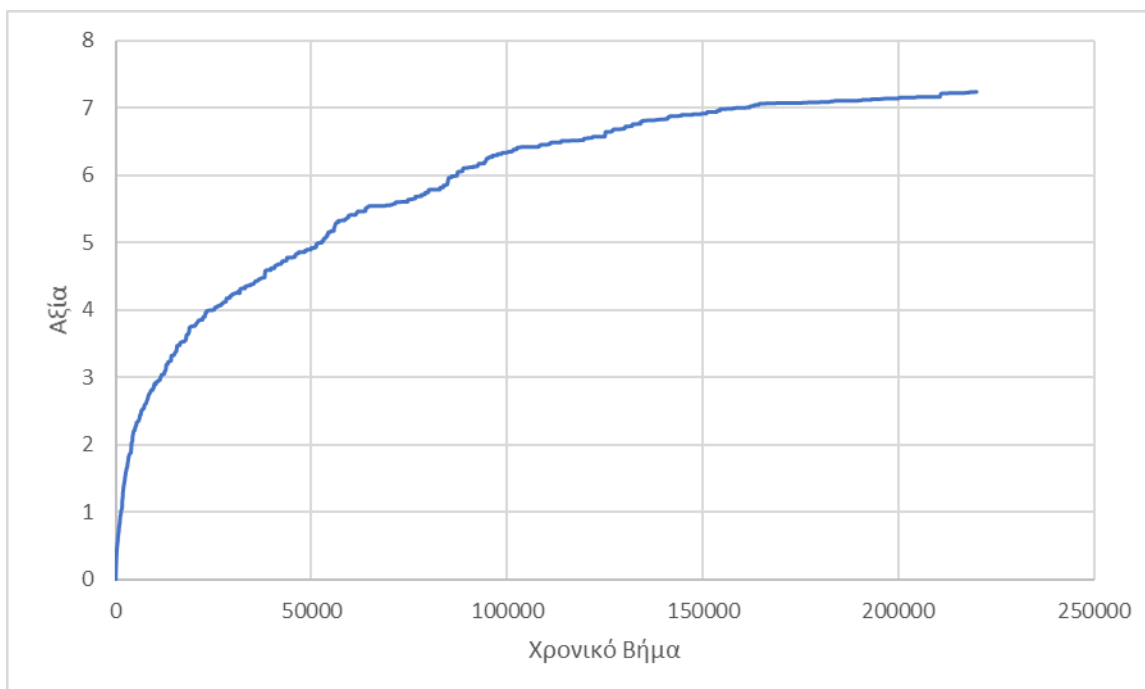
Ανάλυση Εκπτώτικου Παράγοντα

Ο εκπτώτικος παράγοντας γ ανταποκρίνεται στο πόσο σημαντικές θεωρούνται οι μελλοντικές ανταμοιβές. Μικρές τιμές του γ δίνουν μεγαλύτερο βάρος στην τωρινή ανταμοιβή, ενώ μεγάλες τιμές ρίχνουν το βάρος στην ανταμοιβή της επόμενης κατάστασης. Στα Διαγράμματα 11 και 12 παρουσιάζεται η διαχρονική μεταβολή της μέσης τιμής των βέλτιστων ενεργειών όλων των καταστάσεων για τιμές της μεταβλητής γ ίσες με 0.7 και 0.9 αντίστοιχα σύμφωνα με τα σενάρια 6 και 3.



Διάγραμμα 11: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $\gamma=0.7$

Γίνεται εμφανές ότι μεγαλύτερες τιμές του γ οδηγούν σε μεγαλύτερες αξίες, το οποίο συμβάλει στον καλύτερο διαχωρισμό των ενεργειών και στην ευκολότερη επιλογή από πλευράς του πράκτορα. Οι μεγαλύτερες τιμές αυτές δεν διαφοροποιούν την ποιότητα του αποτελέσματος αυτή καθαυτή καθώς δεν σημαίνει ότι μεγαλύτερες τιμές συνεπάγονται καλύτερα αποτελέσματα. Αλλά διαχωρίζουν με πιο έντονο τρόπο τις επιθυμητές από τις ανεπιθύμητες ενέργειες. Αυτό είναι ιδιαίτερα ωφέλιμο σε περιπτώσεις όπου το περιβάλλον είναι πολύπλοκο και αποτελείται από συνεχείς μεταβλητές.

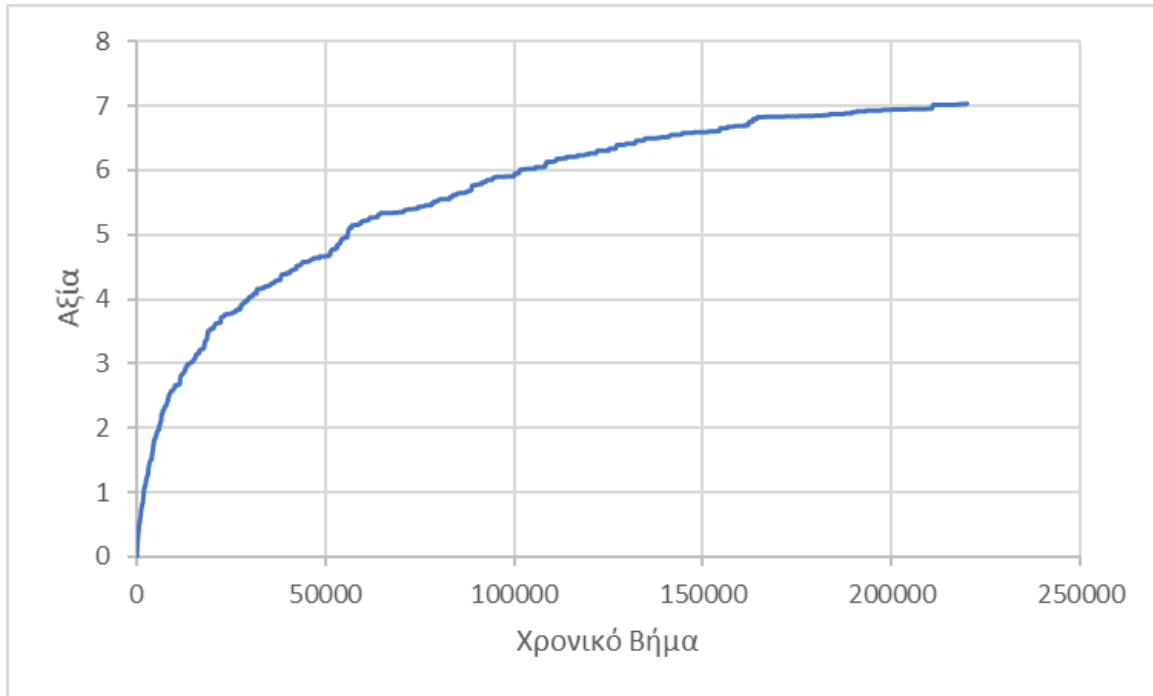


Διάγραμμα 12: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $\gamma=0.9$

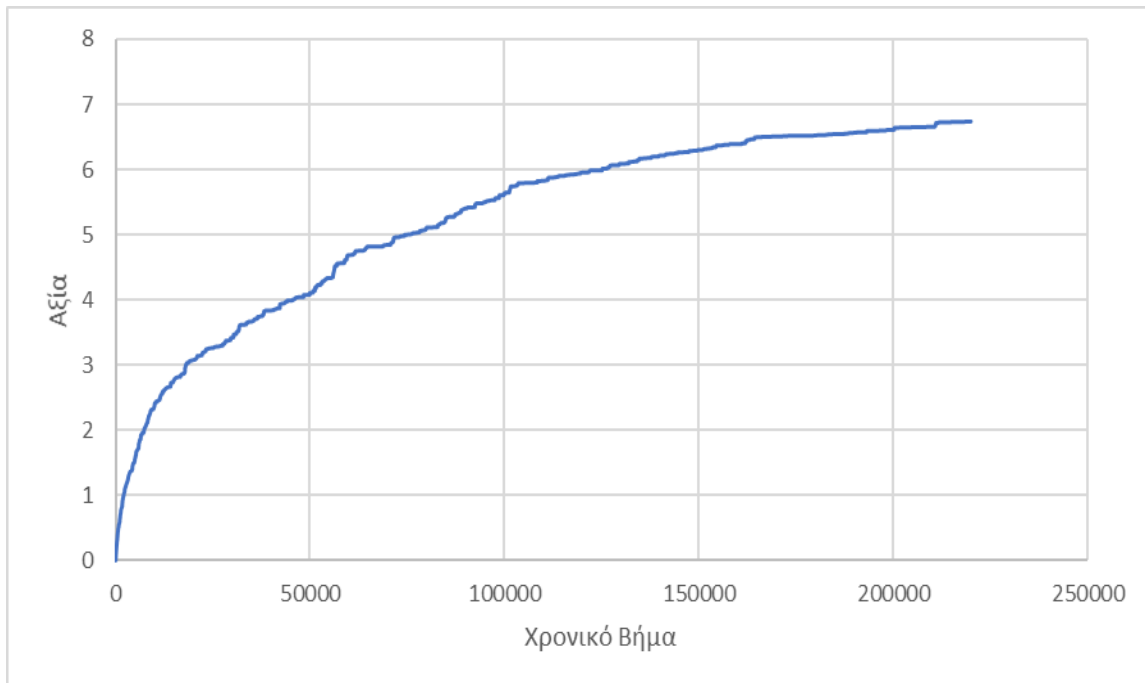
Ανάλυση Βαθμού Εξερεύνησης

Ο βαθμός εξερεύνησης αντιπροσωπεύει το ποσοστό των χρονικών βημάτων κατά το οποίο ο πράκτορας επιλέγει τυχαία ενέργεια. Έτσι, με την χρήση μιας συνάρτησης παραγωγής τυχαίων αριθμών επιλέγεται ενέργεια αγνοώντας την πολιτική. Τα Διαγράμματα 13 και 14 παρουσιάζουν για τα σενάρια 4 και 5 αντίστοιχα τις επιδράσεις των διαφορετικών τιμών του ϵ στην μέση αξία όλων των βέλτιστων ενεργειών των καταστάσεων. Τα διαγράμματα είναι παρόμοια για τις δύο τιμές της παραμέτρου εξερεύνησης. Αυτό μπορεί να οφείλεται στο γεγονός ότι το πρόβλημα είναι ρητά ορισμένο και δομημένο με αρκετά ξεκάθαρο τρόπο έτσι ώστε να μην αποτελεί πρόβλημα η σύγκλιση του αλγορίθμου ενώ υπάρχουν αρκετά δεδομένα για εκπαίδευση. Ωστόσο σε προβλήματα όπου το περιβάλλον είναι πολύπλοκο και ο τρόπος που έχει επιλεγεί για την αντιμετώπιση του προβλήματος δεν είναι σαφώς ορισμένος η εξερεύνηση παραμέτρων είναι αναγκαία για την εύρεση του ιδανικού συνδυασμού.

5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ



Διάγραμμα 13: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $e=0.5$



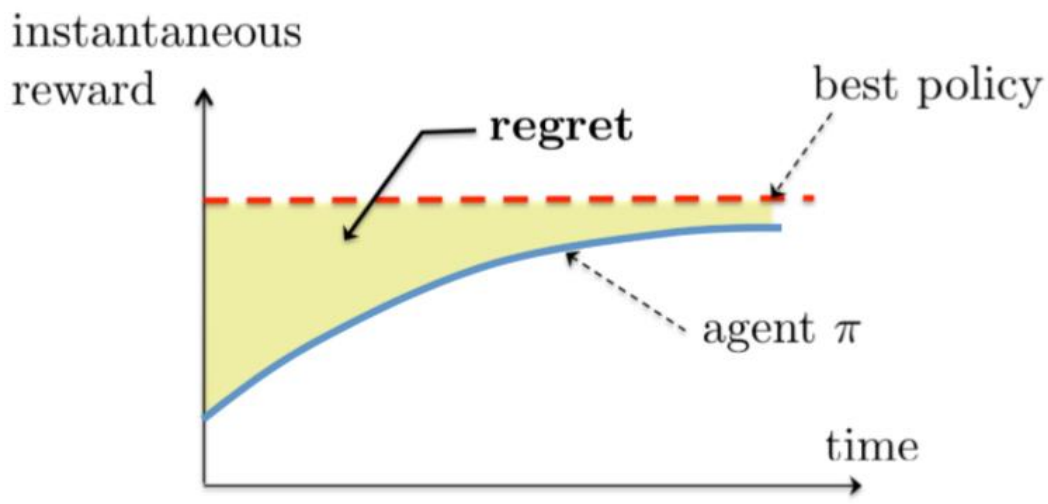
Διάγραμμα 14: Αξία της μέσης βέλτιστης ενέργειας των καταστάσεων, $e=0.7$

Μετανόηση

Ένας επιπλέον τρόπος για την αξιολόγηση των αποτελεσμάτων είναι η σύγκριση της μετανόησης (**regret**) των αποτελεσμάτων για τους διάφορους συνδυασμούς παραμέτρων. Η μετανόηση ενός αλγορίθμου π είναι η διαφορά της αθροιστικής ανταμοιβής της βέλτιστης πολιτικής από την αθροιστική ανταμοιβή της πολιτικής π που ακολουθήθηκε.

Η μετανόηση της πολιτικής π μετά από T χρονικά βήματα υπολογίζεται ως

$$R^\pi(T) = TV_H^{\pi^*} - \sum_{t=1}^T \mathbb{E}[V_H^{\pi_t}] \quad (2)$$



Εικόνα 13: Έννοια της μετανόησης (Πηγή: Proutiere et al. 2019)

Ο υπολογισμός της μετανόησης γίνεται μετά από 220000 χρονικά βήματα για κάθε σενάριο, την στιγμή που τελειώνει η εκπαίδευση. Η μέγιστη αθροιστική ανταμοιβή που μπορεί να λάβει ο αλγόριθμος σε κάθε σενάριο είναι 220000 καθώς η μέγιστη ανταμοιβή για κάθε χρονικό βήμα είναι 1 σύμφωνα με την συνάρτηση ανταμοιβής. Οι τελικές αθροιστικές ανταμοιβές κάθε σεναρίου είναι οι εξής:

1. 43233.5
2. 43175
3. 131490.5
4. -425.5
5. -87961
6. 131507

Οπότε η μετανόηση για το κάθε σενάριο είναι:

1. 156766.5
2. 156825

3. 68509.5
4. 200425.5
5. 287961
6. 68493

Τα σενάρια 4 και 5 παρουσιάζουν ιδιαίτερα μεγάλες τιμές αυτό οφείλεται στο ότι έχουν μεγάλο ποσοστό εξερεύνησης οπότε επιλέγονται τυχαία αρκετές από τις ενέργειες που λαμβάνει ο πράκτορας και όχι με βάση την πολιτική. Τα σενάρια 3 και 6 εμφανίζουν την ιδανικότερη μετανόηση, καθώς ο βαθμός εξερεύνησης είναι μικρός οπότε ο αλγόριθμος ακολουθεί την πολιτική που μαθαίνει κατά κύριο λόγο και δεν εξερευνά καινούργιες ενέργειες.

Ακρίβεια

Ένας τελευταίος τρόπος για να αξιολογηθούν τα διαφορετικά σενάρια είναι η ακρίβεια των αποτελεσμάτων. Για τον υπολογισμό της ακρίβειας ο εκπαιδευμένος αλγόριθμος καλείται να επιλέξει ενέργειες για το σύνολο των δεδομένων αξιολόγησης και να λάβει ανταμοιβές ανάλογα με τις ενέργειες που επέλεξε. Τα δεδομένα αξιολόγησης αποτελούνται από 54976 γραμμές, δηλαδή πραγματοποιήθηκαν 54976 επαναλήψεις για κάθε σενάριο. Υπολογίζοντας σε πόσες από αυτές τις επαναλήψεις ο πράκτορας επέλεξε λάθος ενέργεια μπορούν να εξαχθούν συμπεράσματα σχετικά με την ποιότητα του αλγορίθμου. Οι λανθασμένες επιλογές για κάθε σενάριο είναι:

1. 10
2. 13
3. 4
4. 6
5. 10
6. 9

Γενικά όλα τα σενάρια παρουσιάζουν ελάχιστο αριθμό λαθών, δεδομένου ότι εξετάστηκαν σε 54976 χρονικές στιγμές. Αυτό συμβαίνει διότι, ασχέτως των παραμέτρων, όλα τα σενάρια είχαν την ευκαιρία να εκτεθούν σε έναν μεγάλο αριθμό καταστάσεων και να μάθουν από αυτές. Βέβαια, ένας αλγόριθμος ο οποίος προορίζεται για μηνύματα σε οδηγούς κατά την διάρκεια της οδήγησης δεν θα μπορούσε να γίνει δεκτός με έναν σχετικά μεγάλο αριθμό λαθών. Έτσι ενώ σε άλλα προβλήματα ενδεχομένως θα γινόταν δεκτή μικρότερη ακρίβεια στην συγκεκριμένη περίπτωση τα λάθη πρέπει να είναι ελάχιστα.

Τελικός Αλγόριθμος

Ο τελικός αλγόριθμος αξιοποίησε τις παραμέτρους του σεναρίου 3 καθώς αυτός ο συνδυασμός παραμέτρων εμφανίζει τα λιγότερα λάθη, μικρή μετανόηση και συγκλίνει. Η τιμή του $\gamma=0.9$ δείχνει ότι ο αλγόριθμος ενδιαφέρεται αρκετά για τις μελλοντικές ανταμοιβές ενώ το $\epsilon=0.2$ επιτρέπει στον αλγόριθμο να εξερευνήσει σε ικανοποιητικό βαθμό χωρίς να οδηγεί σε τυχαία επιλογή ενεργειών για μεγάλο αριθμό επαναλήψεων και τέλος ο ρυθμός εκμάθησης $lr=0.2$ αποτελεί μια αρκετά τυπική τιμή έτσι ώστε να μην αγνοείται η ήδη αποκτηθείσα γνώση για χάρη της καινούργιας γνώσης.

Με βάση τα παραπάνω προκύπτει ο **βέλτιστος Πίνακας-Q** (Πίνακας 3), ο οποίος περιλαμβάνει τις **τιμές των αξιών για όλους τους συνδυασμούς καταστάσεων-ενεργειών**. Μέσω του πίνακα είναι δυνατό να επιλεχθεί η βέλτιστη ενέργεια για κάθε κατάσταση στην οποία μπορεί να βρεθεί το περιβάλλον. Παρατηρείται ότι ορισμένες καταστάσεις έχουν μηδενικές αξίες για όλες τις ενέργειες. Αυτό οφείλεται στο γεγονός ότι ο αλγόριθμος δεν εκτέθηκε σε αυτές τις καταστάσεις κατά την διάρκεια της εκπαίδευσης. Επίσης ορισμένες μηδενικές αξίες σε ενέργειες οφείλονται στο γεγονός ότι ο αλγόριθμος δεν εξερεύνησε επαρκώς αυτές τις καταστάσεις διότι πιθανώς δεν τις συνάντησε αρκετές φορές ώστε να δοκιμάσει και άλλες ενέργειες.

Ο τελικός πίνακας παρουσιάζει ελλείψεις κατά κύριο λόγο σε ορισμένες ακραίες περιπτώσεις. Το πρόβλημα αυτό μπορεί να αποδοθεί στο γεγονός ότι παρόλο που το δείγμα με βάση το οποίο έμαθε να ενεργεί ο αλγόριθμος είναι μεγάλο, δεν περιέχει πολλές ακραίες περιπτώσεις. Κατά την διάρκεια των μετρήσεων δεν παρατηρήθηκαν συγκρούσεις ή οριακές συγκρούσεις έτσι ώστε ο αλγόριθμος να μάθει να συμπεριφέρεται σε ιδιαίτερα επικίνδυνες καταστάσεις.

Ο πίνακας αυτός μπορεί να χρησιμοποιηθεί στα πλαίσια ενός **συστήματος υποβοήθησης οδηγού**, όπου θα δίνει βοηθητικά μηνύματα κατά την διάρκεια της οδήγησης για την αποφυγή συγκρούσεων με άλλα οχήματα της κυκλοφορίας. Στο πλαίσιο της μετάβασης σε αυτόματη κυκλοφορία, ο αλγόριθμος θα μπορούσε να αποτελέσει **μέρος του συστήματος αυτόνομης οδήγησης**, καθώς είναι σε θέση να διαχειριστεί την αλληλεπίδραση του οχήματος με την γύρω κυκλοφορία, να αναγνωρίσει τυχόν επικίνδυνες καταστάσεις και να επιλέξει τις κατάλληλες ενέργειες.

Πίνακας 3: Πίνακας-Q

Action State	1	2	3	4	5	6	7	8	9
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	10.000	8.485	8.489	8.387	8.488	6.963	6.982	6.953	6.927
3	1.620	1.080	1.021	10.000	2.430	4.699	0.000	4.706	3.347
4	0.698	1.807	-0.400	1.378	9.313	0.000	0.000	-0.100	0.000
5	3.457	0.000	0.000	0.137	0.263	0.000	0.000	0.000	0.000
6	-0.100	-0.400	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.331	-0.653	-0.400	-0.100	3.177	-0.206	0.000	0.425	0.000
8	1.829	0.000	0.105	0.000	0.000	0.000	0.000	-0.329	0.000
9	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
10	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
11	5.681	4.668	2.264	10.000	5.715	3.057	5.018	5.531	6.135
12	7.335	5.448	5.060	5.427	10.000	6.615	6.247	6.717	8.005
13	10.000	8.260	8.191	8.433	8.420	6.936	6.913	6.948	6.942
14	1.620	-0.400	-0.400	10.000	1.045	1.700	1.681	2.347	0.000
15	1.620	1.748	1.080	-0.100	10.000	0.000	1.262	0.000	0.000
16	6.358	1.281	0.000	0.000	1.392	0.964	0.978	0.457	0.000
17	1.609	10.000	1.398	3.222	2.467	3.053	4.026	0.000	2.487
18	2.996	10.000	1.600	0.000	3.408	1.505	3.268	2.903	1.300
19	-0.400	-0.100	-0.720	0.000	0.000	0.000	0.000	0.000	0.000
20	-0.400	-0.100	0.000	-0.400	0.000	0.000	0.000	0.000	0.000
21	-0.100	6.703	-0.100	-0.400	0.658	1.106	0.000	0.000	0.000
22	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
23	-0.400	0.000	0.000	-0.400	0.000	0.000	-0.400	0.000	0.000
24	-0.100	1.038	0.000	-0.400	0.000	0.000	0.000	0.000	0.000
25	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
26	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
27	8.223	6.763	6.865	10.000	8.306	7.587	8.259	6.060	6.966
28	7.761	5.448	6.740	8.026	10.000	6.678	6.948	8.022	8.261
29	9.999	6.833	6.968	7.823	7.265	6.632	6.585	4.556	4.927
30	1.592	1.878	3.018	9.998	4.233	1.699	2.944	1.270	3.963
31	4.298	4.925	5.811	3.410	9.998	2.755	4.168	5.803	5.696
32	9.346	0.006	4.166	0.000	2.304	2.806	3.265	1.070	4.097
33	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
34	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
35	8.500	7.000	7.000	10.000	8.500	8.500	8.500	7.000	7.000
36	8.500	7.000	7.000	8.500	10.000	7.000	7.000	8.500	8.500

5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

Action State	1	2	3	4	5	6	7	8	9
37	10.000	8.500	8.500	8.500	8.500	7.000	7.000	7.000	7.000
38	6.246	6.638	6.867	10.000	7.580	8.448	8.029	6.842	6.517
39	8.500	6.967	6.919	8.474	9.992	6.991	6.977	8.460	8.377
40	10.000	8.165	8.306	8.281	8.111	6.873	6.473	6.630	6.303
41	7.056	10.000	7.916	6.754	6.687	7.587	6.007	8.201	6.842
42	8.115	10.000	8.201	6.614	6.615	8.033	6.898	8.450	6.842
43	-0.400	2.996	0.000	3.057	2.264	10.000	5.106	3.953	3.352
44	-0.400	-0.100	1.080	-0.400	2.792	1.318	2.264	10.000	0.000
45	-0.100	10.000	0.000	4.115	3.415	1.456	0.000	4.135	1.400
46	0.887	-0.100	-0.400	-0.100	-0.400	9.141	0.000	0.963	0.000
47	0.783	-0.100	-0.400	0.158	-0.100	1.465	0.490	9.970	0.000
48	-0.100	9.636	-0.100	0.917	1.272	0.000	-0.400	0.000	0.000
49	3.938	5.681	10.000	4.920	3.352	5.356	5.714	3.416	4.148
50	4.097	1.620	10.000	3.290	3.373	2.516	4.919	4.993	1.647
51	-0.400	1.481	1.427	-0.100	0.887	-0.100	8.926	0.248	0.000
52	1.781	-0.400	-0.100	1.040	1.318	-0.400	-0.100	0.468	9.648
53	2.033	0.000	8.530	0.000	0.000	0.000	0.000	0.000	0.000
54	-0.400	-0.400	-0.180	-0.100	-0.400	-0.100	7.867	0.000	0.000
55	-0.400	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
56	-0.100	-0.100	4.009	0.000	0.000	0.000	0.000	0.000	0.000
57	7.346	5.682	10.000	6.918	6.781	5.826	8.197	6.840	8.122
58	8.027	7.570	10.000	6.466	6.612	6.322	8.027	6.692	6.950
59	2.996	-0.400	2.259	1.398	1.098	3.744	10.000	1.399	1.700
60	1.077	-0.400	2.995	1.111	-0.100	3.967	2.074	-0.100	10.000
61	1.535	4.096	10.000	1.400	0.000	2.499	0.000	0.000	1.228
62	-0.400	-0.400	0.692	1.697	-0.400	-0.100	9.141	-0.400	1.243
63	0.887	-0.400	-0.100	0.608	-0.100	-0.720	-0.100	-0.100	9.450
64	-0.100	1.110	9.928	0.000	0.787	-0.149	0.963	1.217	0.000
65	0.776	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
66	3.324	0.000	0.000	0.000	0.000	0.000	-0.294	0.000	0.000
67	-0.100	0.000	0.000	0.000	0.000	0.000	0.000	-0.400	0.000
68	-0.100	-0.400	-0.400	-0.100	0.000	0.000	0.000	0.000	0.000
69	0.396	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
70	-0.100	-0.400	0.000	0.000	0.000	0.000	0.000	0.000	0.000
71	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.100	0.000
72	0.200	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

6. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ

6.1 ΕΙΣΑΓΩΓΗ

Η οδική ασφάλεια αποτελεί αντικείμενο μείζονος σημασίας και χρήζει ιδιαίτερης προσοχής. Συγκεκριμένα, θα πρέπει να μειωθεί η συμμετοχή του ανθρώπινου παράγοντα από την διαδικασία της οδήγησης, καθώς ο άνθρωπος είναι η κύρια αιτία πρόκλησης ατυχημάτων. Αυτό θα επιτευχθεί με την εισαγωγή συστημάτων υποβοήθησης οδηγού στα οχήματα. Το αντικείμενο της παρούσας Διπλωματικής Εργασίας είναι **η δημιουργία ενός αλγορίθμου ενισχυτικής μάθησης για τη βελτίωση της οδηγικής συμπεριφοράς**. Αυτός ο αλγόριθμος θα πρέπει να εντοπίζει τυχόν επικίνδυνες καταστάσεις και να εμφανίζει στον χρήστη τις κατάλληλες ενέργειες, οι οποίες θα τον οδηγήσουν σε μια ασφαλέστερη κατάσταση.

Τα στοιχεία τα οποία χρησιμοποιήθηκαν για την εκπαίδευση του αλγορίθμου πάρθηκαν από το πείραμα που πραγματοποιήθηκε σε μια ευρύτερη περιοχή στο κέντρο της Αθήνας, όπου καταγράφηκαν οι τροχιές των οχημάτων με χρήση ΣμηΕΑ. Στη συνέχεια ακολούθησε η επεξεργασία των δεδομένων σε περιβάλλον python και τέλος διαμορφώθηκε ο αλγόριθμος ενισχυτικής μάθησης και έγιναν δοκιμές για την εύρεση του βέλτιστου συνδυασμού παραμέτρων.

Μετά από μια σειρά δοκιμών, αναπτύχθηκε ο βέλτιστος Πίνακας-Q (Πίνακας 3) ο οποίος περιέχει την αξία όλων των συνδυασμών καταστάσεων-ενεργειών. Μέσω αυτού του πίνακα ο πράκτορας είναι σε θέση να επιλέγει σε κάθε χρονική στιγμή την βέλτιστη ενέργεια ελέγχοντας στον πίνακα όλες τις αξίες των ενεργειών δεδομένης της κατάστασης και επιλέγοντας αυτήν με την μεγαλύτερη αξία.

Ο τελικός Πίνακας-Q που προέκυψε από τον αλγόριθμο, μπορεί να χρησιμοποιηθεί για τη βελτίωση της οδηγικής συμπεριφοράς.

6.2 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Στο παρόν υποκεφάλαιο συνοψίζονται τα συμπεράσματα που προέκυψαν από τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας, μέσω της σύνθεσης των αποτελεσμάτων των προηγούμενων κεφαλαίων.

Ο αλγόριθμος έχει την ικανότητα να μαθαίνει γρήγορα όπως προκύπτει και από το διάγραμμα της αξίας της μέσης μέγιστης ενέργειας όλων των καταστάσεων (Διάγραμμα 11) το οποίο συγκλίνει. Ένας αλγόριθμος ενισχυτικής μάθησης, για να γίνει αποδεκτός και να χρησιμοποιηθεί στην πράξη, θα πρέπει να φτάσει σε σύγκλιση κατά την διάρκεια της εκπαίδευσης. Η ταχύτητα με την οποία φτάνει στην σύγκλιση αποτελεί ένδειξη της ποιότητάς του. Συνεπώς ο αλγόριθμος που αναπτύχθηκε στα πλαίσια αυτής της Διπλωματικής Εργασίας μπορεί να θεωρηθεί ικανοποιητικός καθώς συγκλίνει αρκετά γρήγορα.

Όλοι οι αλγόριθμοι ενισχυτικής μάθησης έχουν την ικανότητα να επεκταθούν και να συνεχίσουν να μαθαίνουν χωρίς να υπάρχει ανάγκη δημιουργίας ενός νέου αλγορίθμου

από την αρχή σε αντίθεση με τα στατιστικά μοντέλα. Έτσι λοιπόν, ο **αλγόριθμος που αναπτύχθηκε μπορεί να εκπαιδευτεί περαιτέρω** για να είναι σε θέση να αντιμετωπίζει με καλύτερο τρόπο τις κρίσιμες καταστάσεις. Οι αλγόριθμοι μαθαίνουν ακόμα και κατά την διάρκεια της χρήσης τους αφότου ολοκληρωθεί η εκπαίδευση καθώς εκτίθενται σε καινούργιες καταστάσεις, βελτιώνοντας έτσι τις επιδόσεις τους και την εμπειρία.

Η συγκεκριμένη Διπλωματική Εργασία αποτελεί μια από τις πρώτες έρευνες που **λαμβάνουν υπόψιν και τα πλευρικά οχήματα** κατά την διάρκεια της αξιολόγησης του επιπέδου ασφάλειας της οδικής κατάστασης. Οι περισσότερες μελέτες πάνω στο αντικείμενο των συστημάτων αποφυγής σύγκρουσης εστιάζουν στο μπροστά και στο πίσω όχημα αγνοώντας την επιρροή στην οδική ασφάλεια των πλευρικών οχημάτων. Ωστόσο, τα πλευρικά οχήματα μπορούν να επηρεάσουν τον τρόπο οδήγησης και να συμβάλλουν στην πρόκληση σύγκρουσης, είτε άμεσα με την εμπλοκή των ίδιων στην σύγκρουση, όπου το όχημα συγκρούεται με το πλευρικό, είτε έμμεσα επηρεάζοντας τον οδηγό λόγω της τοποθέτησής τους σε σχέση με το εξεταζόμενο όχημα.

Ο αλγόριθμος είναι σε θέση να **βελτιώσει την οδηγική συμπεριφορά**, καθώς, με βάση τα πραγματικά δεδομένα των οδηγών, το μοντέλο θα απέφυγε κρίσιμες καταστάσεις όπως προκύπτει από την ανταμοιβή που πήρε ο πράκτορας κατά τον έλεγχο του τελικού αλγορίθμου. Έτσι αν τα οχήματα βρίσκονταν σε κρίσιμες καταστάσεις, ο αλγόριθμος θα τις αναγνώριζε και θα επέλεγε ενέργειες οι οποίες θα τις απέτρεπαν.

Η **ενισχυτική μάθηση αποτελεί κατάλληλη μέθοδο για την διαμόρφωση των αλγορίθμων για τα συστήματα υποβοήθησης οδηγού** και τα συστήματα αυτόνομης οδήγησης. Οι μέθοδοι αυτοί είναι σε θέση να διαχειριστούν περιβάλλοντα τα οποία δεν είναι σαφώς ορισμένα και ο αριθμός των καταστάσεων στις οποίες μπορούν να βρεθούν είναι άπειρος καθώς αποτελούνται από πολλές διαστάσεις. Η οδήγηση αποτελεί μια πολύπλοκη διεργασία κατά την οποία πρέπει να ληφθούν υπόψιν διάφοροι παράγοντες για την επιλογή της βέλτιστης ενέργειας. Συνεπώς, οι αλγόριθμοι ενισχυτικής μάθησης είναι σε θέση να αντιμετωπίσουν τέτοιες διεργασίες όπως προκύπτει και από τα αποτελέσματα της παρούσας Διπλωματικής Εργασίας.

Ο αλγόριθμος κατά την διάρκεια του ελέγχου έκανε **4 λάθη στις 54976 δοκιμές**, τα οποία οφείλονται στο ότι ο αλγόριθμος κατά την εκπαίδευση δεν εκτέθηκε σε αυτές τις καταστάσεις επαρκώς έτσι ώστε να εντοπίσει την ιδανική ενέργεια. Τα λάθη είναι πολύ λίγα, καθώς ο αλγόριθμος έχει εκπαιδευτεί με χρήση μιας μεγάλης βάσης δεδομένων, συνεπώς έχει εκτεθεί σε ποικίλες καταστάσεις και έχει μάθει να τις διαχειρίζεται. Ένα σύστημα υποβοήθησης οδηγού θα πρέπει να παρουσιάζει ελάχιστα λάθη, καθώς τυχόν λάθος ενέργειες μπορεί να οδηγήσουν σε σύγκρουση. Η αξιοπιστία τέτοιων συστημάτων πρέπει να είναι υψηλή για να μπορέσουν να γίνουν δεκτά και να χρησιμοποιηθούν.

Ο **τελικός Πίνακας-Q περιέχει μερικά μηδενικά**. Αυτό μπορεί να οφείλεται στο γεγονός ότι κατά την διάρκεια της εκπαίδευσης ο αλγόριθμος δεν εκτέθηκε σε ικανοποιητικό βαθμό σε αυτές τις καταστάσεις έτσι ώστε να δοκιμάσει τις διάφορες ενέργειες μέχρις ότου βρει την βέλτιστη που θα επιφέρει την μέγιστη ανταμοιβή. Ορισμένες καταστάσεις έχουν μηδενικά σε όλες τις ενέργειες. Αυτές οι καταστάσεις είναι αρκετά ακραίες και ο αλγόριθμος δεν είχε την ευκαιρία να εκτεθεί σε πολλές κρίσιμες καταστάσεις. Κατά την διάρκεια συλλογής των δεδομένων δεν παρατηρήθηκαν πολλές κρίσιμες καταστάσεις ή οριακές συγκρούσεις για να μάθει ο αλγόριθμος από αυτές και να επεκτείνει το εύρος των

καταστάσεων το οποίο μπορεί να διαχειριστεί. Συνεπώς ο Πίνακας-Q περιέχει μηδενικά σε αυτές τις καταστάσεις και ο αλγόριθμος δεν θα ήταν σε θέση να επιλέξει την σωστή ενέργεια σε περίπτωση που βρισκόταν σε αυτές τις καταστάσεις.

6.3 ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑ

Για την περαιτέρω διερεύνηση του αντικειμένου της παρούσας Διπλωματικής Εργασίας, ενδιαφέρον θα παρουσίαζε να διερευνηθούν τα παρακάτω:

Χρήση βαθιάς ενισχυτικής μάθησης

Πρόκειται για τον συνδυασμό της ενισχυτικής μάθησης με νευρωνικά δίκτυα σε περιπτώσεις όπου το περιβάλλον είναι πολύπλοκο και δεν είναι δυνατό να αποθηκευτούν όλοι οι συνδυασμοί κατάστασης-ενέργειας. Έτσι, προστίθεται ένα νευρωνικό δίκτυο το οποίο προβλέπει την Q-αξία, με βάση τις μεταβλητές εισόδου (κατάσταση, ενέργεια). Ο όρος “βαθιά” αναφέρεται στην ύπαρξη ορισμένων κρυμμένων επιπέδων για να αποτυπώνονται όλες οι πολύπλοκες λεπτομέρειες του περιβάλλοντος. Με την βαθιά ενισχυτική μάθηση θα ήταν δυνατό να διαμορφωθεί το περιβάλλον χρησιμοποιώντας όλο το εύρος των μεγεθών που το περιγράφουν, χωρίς να υπάρχει η ανάγκη οριοθέτησης των διαστημάτων.

Κατηγοριοποίηση οδηγών της γύρω κυκλοφορίας

Όσο ένα όχημα βρίσκεται εντός του πεδίου του εξεταζόμενου οχήματος, ο αλγόριθμος καταγράφοντας τις ταχύτητες και τις επιταχύνσεις του, να το κατηγοριοποιεί ως προς την επιθετικότητα του. Όταν ένας οδηγός είναι επιθετικός και οδηγεί επικίνδυνα λοιπόν, ο αλγόριθμος να είναι επιφυλακτικός ως προς τον οδηγό αυτό και να προτείνει πιο συντηρητικές ενέργειες στον οδηγό του εξεταζόμενου οχήματος, λαμβάνοντας υπόψιν την αβεβαιότητα των γύρω οχημάτων. Με αυτόν τον τρόπο οι ενέργειες που θα προτείνει ο αλγόριθμος θα είναι πιο συγκεκριμένες και θα μειώνονται οι πιθανότητες σύγκρουσης.

Συμπεριφορά και χαρακτηριστικά του οδηγού

Ο χρόνος αντίδρασης, ο τρόπος οδήγησης, η επιλογή ταχυτήτων κλπ, διαφέρουν σε κάθε οδηγό. Έτσι, ο αλγόριθμος λαμβάνοντας υπόψιν τα χαρακτηριστικά του κάθε οδηγού ξεχωριστά θα είναι σε θέση να προτείνει εξατομικευμένες ενέργειες και να βελτιώνει τα επίπεδα της οδικής ασφάλειας.

Εκπαίδευση με χρήση περισσότερων κρίσιμων καταστάσεων

Ο αλγόριθμος θα πρέπει να εκτεθεί σε περισσότερες κρίσιμες καταστάσεις έτσι ώστε να μάθει να τις διαχειρίζεται με τον βέλτιστο δυνατό τρόπο. Αυτό μπορεί να πραγματοποιηθεί, είτε με την χρήση περισσότερων πραγματικών δεδομένων εκπαίδευσης, τα οποία θα περιέχουν προφανώς περισσότερες κρίσιμες καταστάσεις, είτε με την απομόνωση ακραίων καταστάσεων και την εισαγωγή αυτών στον αλγόριθμο. Η εύρεση των απομονωμένων κρίσιμων καταστάσεων μπορεί να γίνει με καταγραφή σε προστατευόμενο περιβάλλον επικίνδυνων οδηγικών συμπεριφορών.

Εισαγωγή περισσότερων μεταβλητών όπως καιρικές συνθήκες, συνθήκες ορατότητας και κυκλοφοριακοί φόρτοι

Οι καιρικές συνθήκες μπορούν να επηρεάσουν το απαιτούμενο χρόνο πέδησης λόγω μειωμένης τριβής του οδοστρώματος. Οπότε, όταν βρέχει θα πρέπει ο αλγόριθμος να προσαρμόζει τους χρόνους πέδησης με βάση την διαθέσιμη τριβή του οδοστρώματος και να προτείνει ανάλογες ενέργειες. Επίσης σε περιπτώσεις χαμηλής ορατότητας η οδήγηση θα πρέπει να είναι πιο συντηρητική και οι αποστάσεις που διατηρούνται μεγαλύτερες για να μειωθεί η πιθανότητα πρόκλησης ατυχήματος. Τέλος οι κυκλοφοριακοί φόρτοι μπορούν να χρησιμοποιηθούν για τον καθορισμό του είδους τις οδήγησης και των αποδεκτών αποστάσεων. Έτσι σε υψηλούς κυκλοφοριακούς φόρτους ο αλγόριθμος να προσαρμόζει τις επιτρεπτές πλευρικές αποστάσεις και τους αποδεκτούς χρόνους για σύγκρουση για να ανταποκρίνονται στην γύρω κυκλοφορία. Αντιθέτως, σε περιπτώσεις χαμηλών κυκλοφοριακών φόρτων, όπου η ροή είναι ελεύθερη και ο οδηγός δεν αναμένει να φρενάρει και να μεταβάλει την ταχύτητα του συνέχεια, οι τιμές των μεγεθών θα πρέπει να είναι πιο συντηρητικές και να λαμβάνουν υπόψιν αυξημένο χρόνο αντίδρασης.

BIBΛΙΟΓΡΑΦΙΑ

World Health Organization (WHO). (2019). Global status report on road safety 2018 (2018).

Γιαννής, Γ. (2018) Τροχαία Ατυχήματα Οδική Συμπεριφορά και Ασφάλεια. Retrieved October 11, 2019, from <https://www.nrso.ntua.gr/geyannis/wp-content/uploads/geyannis-cp329.pdf>

Φραντζεσκάκης, Ι. Μ., Ι. Κ. Γκόλιας (1994) “Οδική Ασφάλεια” Εκδόσεις Παπασωτηρίου, Αθήνα.

Cooper, G. (2019). The Evolution of Deep Learning for ADAS Applications. Retrieved October 11, 2019 from <https://www.synopsys.com/designware-ip/technical-bulletin/deep-learning-dwtb-q217.html>

Papathanasopoulou, V., & Antoniou, C. (2015). Towards data-driven car-following models. *Transportation Research Part C: Emerging Technologies*, 55, 496-509.

Liao, P., Tang, T. Q., Wang, T., & Zhang, J. (2019). A car-following model accounting for the driving habits. *Physica A: Statistical Mechanics and its Applications*, 525, 108-118.

Fu, R., Li, Z., Sun, Q., & Wang, C. (2019). Human-like car-following model for autonomous vehicles considering the cut-in behavior of other vehicles in mixed traffic. *Accident Analysis & Prevention*, 132, 105260.

Yu, A., Palefsky-Smith, R., & Bedi, R. (2016). Deep reinforcement learning for simulated autonomous vehicle control. *Course Project Reports: Winter*, 1-7.

Zhu, M., Wang, X., & Wang, Y. (2018). Human-like autonomous car-following model with deep reinforcement learning. *Transportation research part C: emerging technologies*, 97, 348-368.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529.

Varone, M., Mayer, D., Melegari, A., & Scagliarini, L. (2017). What is machine learning? a definition Retrieved October 11, 2019, from <https://www.expertsystem.com/machine-learning-definition/>

Schulman, J. (2016). Optimizing expectations: From deep reinforcement learning to stochastic computation graphs (Doctoral dissertation, UC Berkeley).

Ashraf, M. (2018). Reinforcement Learning Demystified: A Gentle Introduction. *Towards Data Science*. Apr, 7.

Ashraf, M. (2018). Reinforcement Learning Demystified: Markov Decision Processes (Part 1). Retrieved October 11, 2019, from <https://towardsdatascience.com/reinforcement-learning-demystified-markov-decision-processes-part-1-bf00dda41690>

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

- Rastogi, D. (2017). Deep Reinforcement Learning for Bipedal Robots.
- Choudhary, A. (2019). A Hands-On Introduction to Deep Q-Learning using OpenAI Gym in Python. Retrieved October 11, 2019, from <https://www.analyticsvidhya.com/blog/2019/04/introduction-deep-q-learning-python/>
- Silver, D. (2015). Lecture 1: Introduction to reinforcement learning. UCL. Retrieved from <http://www0.cs.ucl.ac.uk/staff/D.Silver/web/Teaching.html>
- Silver, D. (2015). Lecture 2: Markov Decision Processes. UCL. Retrieved from www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching_files/MDP.pdf
- Barmponakis, E., & Geroliminis, N. (2019). Utilizing a swarm of drones for large-scale traffic measurements.
- Barmponakis, E. N., Vlahogianni, E. I., & Golias, J. C. (2016). Unmanned Aerial Aircraft Systems for transportation engineering: Current practice and future challenges. *International Journal of Transportation Science and Technology*, 5(3), 111-122.
- Barmponakis, E. N., Vlahogianni, E. I., Golias, J. C., & Babinec, A. (2019). How accurate are small drones for measuring microscopic traffic parameters?. *Transportation Letters*, 11(6), 332-340.
- ΑΝΤΩΝΙΟΥ, Β. (2018). Προβολικά Συστήματα: Γεωγραφικά Συστήματα Πληροφοριών. Retrieved October 11, 2019, from <https://eclass.uoa.gr/modules/document/file.php/GEOL123/Y2204%28μάθημα%29/GIS-Ευελπίδου/ProvolikaSysthmata.pdf>
- Minderhoud, M. M., & Bovy, P. H. (2001). Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention*, 33(1), 89-97.
- Hogema, J. H., & Janssen, W. H. (1996). Effects of intelligent cruise control on driving behaviour: a simulator study (No. 1996 C012). TNO.
- Van der Horst, A. R. A. (1991). A time-based analysis of road user behaviour in normal and critical encounters.
- Arbabzadeh, N., Jafari, M., Jalayer, M., Jiang, S., & Kharbeche, M. (2019). A hybrid approach for identifying factors affecting driver reaction time using naturalistic driving data. *Transportation research part C: emerging technologies*, 100, 107-124.
- Johansson, G., & Rumar, K. (1971). Drivers' brake reaction times. *Human factors*, 13(1), 23-27.
- Magister, T., Krulec, R., Batista, M., & Bogdanovic, L. (2006). Measuring a Driver's Reaction Time. *Strojniski Vestnik*, 52(1), 26-40.
- Lerner, N. D., Huey, R. W., McGee, H. W., & Sullivan, A. (1995). Older Driver Perception-Reaction Time for Intersection Sight Distance and Object Detection. Volume I: Final Report. Report No. FHWA-RD-93-168, Federal Highway Administration.
- Qu, D., Chen, X., Yang, W., & Bian, X. (2014). Modeling of car-following required safe distance based on molecular dynamics. *Mathematical Problems in Engineering*, 2014.

Proutiere, A., Talebi, S., %Ok, J. (n.d.). Reinforcement Learning Lecture 3: RL problems, sample complexity and regret. Retrieved October 11, 2019, from [http://www.it.uu.se/research/systems and control/education/2017/relearn/lec3.pdf](http://www.it.uu.se/research/systems%20and%20control/education/2017/relearn/lec3.pdf)

Galarnyk, M. (2018). Explaining the 68-95-99.7 rule for a Normal Distribution. Retrieved October 11, 2019 from <https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2>