

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

ΣΥΣΤΑΔΟΠΟΙΗΣΗ ΔΕΔΟΜΕΝΩΝ ΚΙΝΗΣΗΣ ΜΕ ΧΡΗΣΗ ΒΑΘΙΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Αγάπη Παρρά

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και
Ασφαλιστικής Επιστήμης του Πανεπιστημίου
Πειραιώς ως μέρος των απαιτήσεων για την
απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς

Δεκέμβριος 2021

Σελίδα έγκρισης

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών.

Τα μέλη της Επιτροπής ήταν:

- Αναπληρωτής Καθηγητής Νίκος Πελέκης (Επιβλέπων)
- Επίκουρος Καθηγητής Χ. Ευαγγελάρας
- Αναπληρωτής Καθηγητής Ελ. Κοφίδης

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS

School of Finance and Statistics



Department of Statistics and Insurance Science

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**CLUSTERING MOBILITY DATA WITH
DEEP NEURAL NETWORKS**

Agapi Parra

MSc Dissertation

submitted to the Department of Statistics and
Insurance Science of the University of Piraeus in
partial fulfilment of the requirements for the
degree of Master of Science in Applied Statistics

Piraeus, Greece

December 2021

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Νίκο Πελέκη που μου έδωσε τη δυνατότητα να πραγματοποιήσω την παρούσα εργασία καθώς επίσης και για την γνώση που μου έδωσε, την έμπνευση και την συμβολή του στην ολοκλήρωση της διπλωματικής εργασίας άλλα και κατά τη διάρκεια του Μεταπτυχιακού Προγράμματος.

Περίληψη

Η ταχεία ανάπτυξη των ασύρματων υποδομών και των τεχνολογιών απόκτησης δεδομένων έχει συμβάλει στην εκθετική αύξηση του όγκου των δεδομένων προς επεξεργασία ειδικά των δεδομένων κίνησης (trajectory data) τα οποία καταγράφουν τις θέσεις κινούμενων αντικειμένων ανά χρονική στιγμή και αποτελεί πλέον ένα σημαντικό μέσο για την μελέτη της ανθρώπινης συμπεριφοράς ή την επίλυση προβλημάτων οδικής κίνησης όπως το κυκλοφοριακό πρόβλημα. Στην μελέτη αυτή βασιστήκαμε σε μια πρόσφατη εργασία που χρησιμοποιεί μεθόδους βαθιάς μάθησης για να κάνει embedding τον αρχικό χώρο των δεδομένων κίνησης και μετά ομαδοποιεί με τον αλγόριθμο K-means. Αφού τον υλοποιήσαμε, τον εξελίξαμε στο κομμάτι της ομαδοποίησης μέσα από 4 αλγορίθμους ομαδοποίησης με διαφορετικά χαρακτηριστικά. Τα αποτελέσματά μας έδειξαν ότι η επιλογή του K-means στην πρωτότυπη εργασία δεν ήταν η βέλτιστη.

Λέξεις κλειδιά: Βαθιά γνώση, Νευρωνικά δίκτυα, Δεδομένα κίνησης

Abstract

The rapid development of wireless infrastructure and data acquisition technologies has contributed to the exponential increase in the volume of data to be processed specifically for trajectory data which records the positions of moving objects by time or period and is now an important means of study of human behavior or solving traffic problems such as traffic problems. In this study we focus on a recent work that uses deep learning methods to embed the original space of motion data and then grouping it with the K-means algorithm. After we implemented it, we developed it in the grouping part through 4 grouping algorithms with different characteristics. Our results showed that the selection of K-means in the original work was not optimal.

Key words: Deep learning, Neural Networks, Trajectory Data

Περιεχόμενα

Ευχαριστίες	i
Περίληψη	ii
Abstract	iii
Περιεχόμενα.....	1
Κατάλογος Σχημάτων	3
Κατάλογος Πινάκων	5
Εισαγωγή	6
Κεφάλαιο 1. Δεδομένα κίνησης.....	7
1.1. Εισαγωγή	7
1.2. Ορισμοί	8
1.3. Εφαρμογές των δεδομένων κίνησης	12
1.4. Αντληση δεδομένων κίνησης.....	14
Κεφάλαιο 2. Νευρωνικά δίκτυα.....	16
2.1. Εισαγωγή	16
2.2. Ιστορική αναδρομή	17
2.3. Παραδείγματα νευρωνικών δικτύων.....	20
2.4. Μοντέρνες τεχνικές νευρωνικών δικτύων	24
2.4.1. Recurrent Neural Networks	25
2.4.2. LSTM.....	26
Κεφάλαιο 3. Αλγόριθμοι Συσταδοποίησης.....	29
3.1 Ο αλγόριθμος k-Means	29
3.2 Φασματική Ομαδοποίηση	29
3.3. Αλγόριθμος DBSCAN	30
3.4 Ιεραρχική συσταδοποίηση	32

Κεφάλαιο 4. Βιβλιογραφική επισκόπηση μεθόδων συσταδοποίησης δεδομένων κίνησης.....	34
Μέρος Α – Παραδοσιακές μέθοδοι.....	34
Μέρος Β – Νεότερες μέθοδοι βασισμένες σε βαθιά μάθηση (DNN).....	41
Κεφάλαιο 5. Πρακτικό Μέρος – Βελτίωση της εργασίας των Yao et al. (2018) – Συγκριτική Αξιολόγηση στο κομμάτι της ομαδοποίησης	70
Συμπεράσματα	75
Παράρτημα - Εγκατάσταση	76
Βιβλιογραφικές αναφορές.....	77

Κατάλογος Σχημάτων

Σχήμα 1 - Κατηγορίες δεδομένων κίνησης (Πηγή: : Lee & Krumm, 2011)	9
Σχήμα 2 - Αποτύπωση 6 διαφορετικών τροχιών στους άξονες XY (Πηγή: Tampakis, et al., 2019)	14
Σχήμα 3 - Βιολογικό και Τεχνητό Νευρωνικό Δίκτυο	16
Σχήμα 4 - Νευρωνικό δίκτυο	21
Σχήμα 5 - Νευρωνικό δίκτυο με ένα κρυφό επίπεδο	22
Σχήμα 6 - Νευρωνικό δίκτυο με δύο κρυφά επίπεδα και περισσότερες από μιας εξόδους	23
Σχήμα 7 - Χρήση νευρωνικού δικτύου για τον υπολογισμό AND, OR και NAND....	24
Σχήμα 8 - Αναπαράσταση των επαναλαμβανόμενων νευρωνικών δικτύων (Α) και των νευρωνικών δικτύων Feedforward (Β). (Πηγή: IBM, 2020)	25
Σχήμα 9 - Μαθηματική αναπαράσταση της μεθόδου RNN (Πηγή: Afshine & Shervine, 2020)	26
Σχήμα 10 - Γραφική απεικόνιση της μεθόδου 2.7. (Πηγή: Balouji et al., 2018).....	27
Σχήμα 11 - Παραδειγμα DBSCAN.....	31
Σχήμα 12 - Αρχικά δεδομένα (αριστερά) και το εξαγόμενο διάγραμμα εγγύτητας (reachability plot με την βοήθεια του αλγόριθμου OPTICS (δεξιά),	35
Σχήμα 13 - Παράδειγμα τροχιάς με 3 σημεία μεταβολής της κίνησης Α,Β και Γ.....	36
Σχήμα 14 - Αρχικά δεδομένα (αριστερά) και συσταδοποίηση υπο-ομάδων (δεξιά)...	36
Σχήμα 15 - Εφαρμογή παραμέτρων συσταδοποίησης με βάση την πυκνότητα (Α. Συντελεστής πυκνότητας =10, Β. Συντελεστής πυκνότητας =31 και Γ. Συντελεστής πυκνότητας = 47).	37
Σχήμα 16 - Αρχικά δεδομένα (Α), Αρχική συσταδοποίηση (Β) και Βέλτιστη συσταδοποίηση (Γ) με την μέθοδο NEAT.....	38
Σχήμα 17 - Αποτελέσματα εφαρμογής αλγορίθμων (α) Πραγματική τροχιά, (b) SFP-L, (c) SFP-C, (d) TPM-L, (e) TPM-C, (f) CACT	39
Σχήμα 18 - Εφαρμογή μεθόδου των Xu et al., (2015), σε δεδομένα ανθρώπινης κίνησης (Α: Αρχικά δεδομένα, Β: Συστάδες, Γ: Route Analysis).....	40
Σχήμα 19 - Σημεία ενδιαφέροντος (α) και εφαρμογή τους σε πραγματικά δεδομένα (β) (Πηγή: Yue et al., 2019).	42
Σχήμα 20 Χωροχρονικό πλέγμα (Πηγή: Harmening et al., 2020).....	46

Σχήμα 21 Δομή ενός νευρωνικού δικτύου αυτόματης κωδικοποίησης. (Πηγή: Olive et al., 2020).	50
Σχήμα 22 Συσταδοποιημένες τροχιές που προσγειώνονται στο αεροδρόμιο της Ζυρίχης, στον διάδρομο 14, εντός 40 ναυτικών μιλίων από το αεροδρόμιο. (Πηγή: Olive et al., 2020).	50
Σχήμα 23 - Εξαγωγή κινητικής συμπεριφοράς. (Πηγή: Yao et al., 2018).	54
Σχήμα 24 - Χρόνος ολίσθησης δημιουργίας παραθύρων. (Πηγή: Yao et al., 2018)	54
Σχήμα 25 - Πλήρης γνωρίσματα. (Πηγή: Yao et al., 2018).	56
Σχήμα 26 - Η δημιουργία ακολουθίας κινούμενης συμπεριφοράς. (Πηγή: Yao et al., 2018)	56
Σχήμα 27 - Αρχιτεκτονική του αυτόματου κωδικοποιητή Seq2Seq. (Πηγή: Yao et al., 2018)	59
Σχήμα 28 - Μέρος των συνθετικών βασικών μοτίβων κίνησης που αποτελούνται από 10 ευθείες τροχιές, 10 κυκλικές τροχιές και 10 καμπύλες. (Πηγή: Yao et al., 2018)	60
Σχήμα 29 - Μέρος του συνθετικού συνδυασμού μοτίβων κίνησης. Κάθε μοτίβο συνδυασμού έχει πέντε τροχιές. Str = Ευθεία, Cir = Κύκλος, Ben = Κάμψη. (Πηγή: Yao et al., 2018)	61
Σχήμα 30 - Το αρχείο καταγραφής μακροπρόθεσμης μνήμης (LSTM) αλλάζει με διαφορετικές παραμέτρους (MSE) . (Πηγή: Yao et al., 2018)	63
Σχήμα 31 - Το αρχείο καταγραφής επαναλαμβανόμενων μονάδων (GRU) αλλάζει με διαφορετικές παραμέτρους (MSE). (Πηγή: Yao et al., 2018)	64
Σχήμα 32 - Χρονοβόρες αλλαγές με διαφορετικές παραμέτρους. (Πηγή: Yao et al., 2018)	65
Σχήμα 33 - Απεικόνιση της μεθόδου ELBOW για να επιλεγθεί η K . \ E_k με την κατάλληλη τιμή K , η οποία σε αυτό το σχήμα πρέπει να είναι το σημείο αγκώνα. Εδώ δίνεται $K = 33$. (Πηγή: Yao et al., 2018)	68
Σχήμα 34 - Απεικόνιση τροχιών στο Cluster 1 (Πηγή: Yao et al., 2018).	68
Σχήμα 35 - Απεικόνιση τροχιών στο Cluster 2 (Πηγή: Yao et al., 2018).	69
Σχήμα 36 - Αξιολόγηση ομαδοποίησης με την μετρική Silhouette.	72
Σχήμα 37 - Αξιολόγηση ομαδοποίησης με την μετρική ARI.	72
Σχήμα 38 - Αξιολόγηση ομαδοποίησης με την μετρική AMI.	73
Σχήμα 39 - Αξιολόγηση ομαδοποίησης με την μετρική Accuracy.	73
Σχήμα 40 - Αξιολόγηση ομαδοποίησης με την μετρική V-measure	74

Κατάλογος Πινάκων

Πίνακας 1 - Ομαδοποίηση της απόδοσης σε συνθετικά δεδομένα. (Πηγή: Yao et al., 2018)	66
Πίνακας 2 - Αποτελέσματα ομαδοποίησης τύπου σκαφών (Πηγή: Yao et al.2018)....	69
Πίνακας 3 - Ακρίβεια αλγορίθμων ανά κατηγορία τροχιών και συνολικά σε ένα πίνακα	71

Εισαγωγή

Η ανάπτυξη στον τομέα της τεχνολογίας της πληροφορίας και της επικοινωνίας, ιδίως στην ανίχνευση θέσης της κινητής τηλεφωνίας και στην ασύρματη επικοινωνία, είχε ως αποτέλεσμα την εύκολη δημιουργία δεδομένων κίνησης, δηλαδή δεδομένα που περιέχουν γεωγραφικές θέσεις στη διάσταση του χρόνου. Η ανάλυσή τους μπορεί να οδηγήσει σε λύσεις για σημαντικά ερευνητικά προβλήματα σε διαφορετικά πεδία όπως πολεοδομικός σχεδιασμός, μεταφορά, οικολογική συμπεριφορά, ανάλυση αθλητικών σκηνών, επιτήρηση και ασφάλεια. Οι ερευνητές στην ανάλυση δεδομένων κίνησης έχουν συμβάλει σημαντικά αναπτύσσοντας μεθόδους και εργαλεία για την επίλυση συγκεκριμένων εφαρμογών και προβλημάτων. Καθώς οι ερευνητές συνεχίζουν να αντιμετωπίζουν διαφορετικά προβλήματα, είναι εμφανής η ανάγκη για την παροχή μιας ολοκληρωμένης εικόνας των προβλημάτων εφαρμογής που επιλύθηκαν χρησιμοποιώντας δεδομένα κίνησης και τις μεθόδους εξόρυξης που είναι διαθέσιμες για τη επίλυση αυτών.

Μια τέτοια ολοκληρωμένη άποψη θα επέτρεπε σε άλλους ερευνητές να εντοπίσουν εύκολα τα ζητήματα που έχουν ήδη επιλυθεί, τα μεθοδολογικά κενά και να αντλήσουν έμπνευση για την ανάπτυξη νέων μεθόδων από τις σχέσεις μεταξύ προβλημάτων και μεθόδων. Ο στόχος αυτής της διατριβής είναι να μελετήσει διάφορους αλγόριθμους ομαδοποίησης τροχιάς και να τους εφαρμόσει σε ένα πραγματικό πρόβλημα που παρουσιάστηκε σε μία πρόσφατη εργασία.

Η δομή της εργασίας είναι η εξής, στο πρώτο κεφάλαιο γίνεται εισαγωγή στα δεδομένα κίνησης. Πιο συγκεκριμένα, γίνεται λεπτομερής αναφορά στους ορισμούς του τύπου των δεδομένων κίνησης στον τρόπο συλλογής τους και στις εφαρμογές τους. Στο δεύτερο κεφάλαιο γίνεται μια εισαγωγή στα νευρωνικά δίκτυα (neural networks) και την βαθιά μάθηση (deep learning). Στο τρίτο κεφάλαιο παρουσιάζονται παραδοσιακοί αλγόριθμοι ομαδοποίησης. Στο τέταρτο κεφάλαιο περιγράφονται οι σχετικές με το θέμα μας εργασίες, τόσο οι πιο καθιερωμένες όσο και οι πιο πρόσφατες που ενσωματώνουν τη γνώση από δίκτυα βαθιάς μάθησης. Στο πέμπτο κεφάλαιο παρουσιάζεται το πρακτικό μέρος της παρούσας διπλωματικής.

Κεφάλαιο 1. Δεδομένα κίνησης

1.1. Εισαγωγή

Ο προσδιορισμός και ο χαρακτηρισμός της κίνησης αντικειμένων είναι σημαντικοί παράγοντες για την υποστήριξη της λήψης αποφάσεων σε τομείς όπως η διαχείριση της ροής της κυκλοφορίας. Ο προσδιορισμός των ροών κίνησης από τροχιές μπορεί να γίνει αυτόματα από δεδομένα τροχιάς εφαρμόζοντας τεχνικές ομαδοποίησης τροχιάς προκειμένου να βρεθεί το βέλτιστο σύνολο συστάδων. Λόγω της σημασίας της στην ανάλυση δεδομένων, η ομαδοποίηση εξακολουθεί να αποτελεί ενεργό αντικείμενο έρευνας. Η εφαρμογή της ομαδοποίησης στις τροχιές είναι ιδιαίτερα δύσκολη για διάφορους λόγους. Πρώτον, ο ορισμός των κατάλληλων μέτρων απόστασης μεταξύ των τροχιών είναι δύσκολος. Δεύτερον, οι τροχιές είναι συχνά διαθέσιμες ως δείγματα σημείων δεδομένων σε χώρο υψηλής διαστάσεων όπου οι κλασικές αποστάσεις χάνουν τη διακριτική τους ισχύ (curse of dimensionality). Τρίτον, η διαθεσιμότητα τεράστιων ποσοτήτων δεδομένων ανοικτής τροχιάς για διερεύνηση απαιτούν αλγόριθμους ομαδοποίησης τροχιάς υψηλής αποτελεσματικότητας και κλιμάκωσης (Xavier et al., 2020).

Οι προσεγγίσεις συσταδοποίησης των δεδομένων τροχιάς που βασίζονται σε κλασικές μεθόδους όπως η k-means ή το DBSCAN παρουσιάζουν παρόμοια προβλήματα. Αυτοί οι καθιερωμένοι αλγόριθμοι εξαρτώνται από συναρτήσεις ομοιότητας (ή απόστασης) που ορίζονται συνήθως απευθείας σε χώρο δεδομένων ανώτερης διάστασης οι οποίες τους εμποδίζουν να συλλάβουν πλουσιότερες εξαρτήσεις που ενδεχομένως βρίσκονται στον λανθάνοντα χώρο χαμηλότερης διάστασης. Για αυτόν τον λόγο, είναι συνήθης πρακτική η χρήση τέτοιων αλγορίθμων να γίνεται με μια τεχνική μείωσης διαστάσεων ως βήμα προ επεξεργασίας. Ωστόσο, εκτός από το γεγονός ότι τεχνικές όπως η ανάλυση πρώτων παραγόντων (PCA) μπορούν να δημιουργήσουν μόνο γραμμικές ενσωματώσεις, αυτή η προσέγγιση των δύο σταδίων μπορεί να είναι μη βέλτιστη καθώς οι παραδοχές στις οποίες βασίζεται η μείωση διαστάσεων και οι τεχνικές ομαδοποίησης είναι ανεξάρτητες. Ο απώτερος στόχος είναι να εκτελείται αυτόματα η εξαγωγή χαρακτηριστικών, η μείωση διαστάσεων και η ομαδοποίηση με ένα και μόνο μοντέλο αντί να εκτελούνται αυτές οι εργασίες ανεξάρτητα με ξεχωριστά μοντέλα. Αυτές οι προσεγγίσεις που βασίζονται στη βαθιά μάθηση (deep learning) έχουν επίσης το

πλεονέκτημα ότι είναι επεκτάσιμες καθώς έχουν αναπτυχθεί για την αντιμετώπιση μεγάλου όγκου δεδομένων γνωστό και ως big data (Xavier et al., 2020).

1.2. Ορισμοί

Καθώς η τεχνολογία της πληροφορίας και επικοινωνίας αναπτύσσεται συνεχώς, τεράστιοι όγκοι δεδομένων προέρχονται από το διαδίκτυο των αντικειμένων(ΔτΑ)¹ Στο ΔτΑ κάθε συσκευή μπορεί να συλλέγει γεωγραφικά δεδομένα (συντεταγμένες) και την χρονική επισήμανση αυτών (time stamp). Αυτό το είδος των μεταβλητών είναι σημαντικές διαστάσεις προς στατιστική ανάλυση και οι συσχετίσεις τους αποτελούν μια σημαντική ιδιότητα ανάκτησης από το ΔτΑ (Chen, Mao & Liu, 2014). Εκτός από το ΔτΑ, και άλλες υπηρεσίες μπορούν να παρέχουν παρόμοια δεδομένα. Κάποιες από αυτές είναι το παγκόσμιο σύστημα εντοπισμού θέσης (GPS), το σύστημα αναγνώρισης ραδιοσυχνότητας (Radio Frequency Identification- RFID), συστήματα κινητής τηλεφωνίας GSM κ.α.

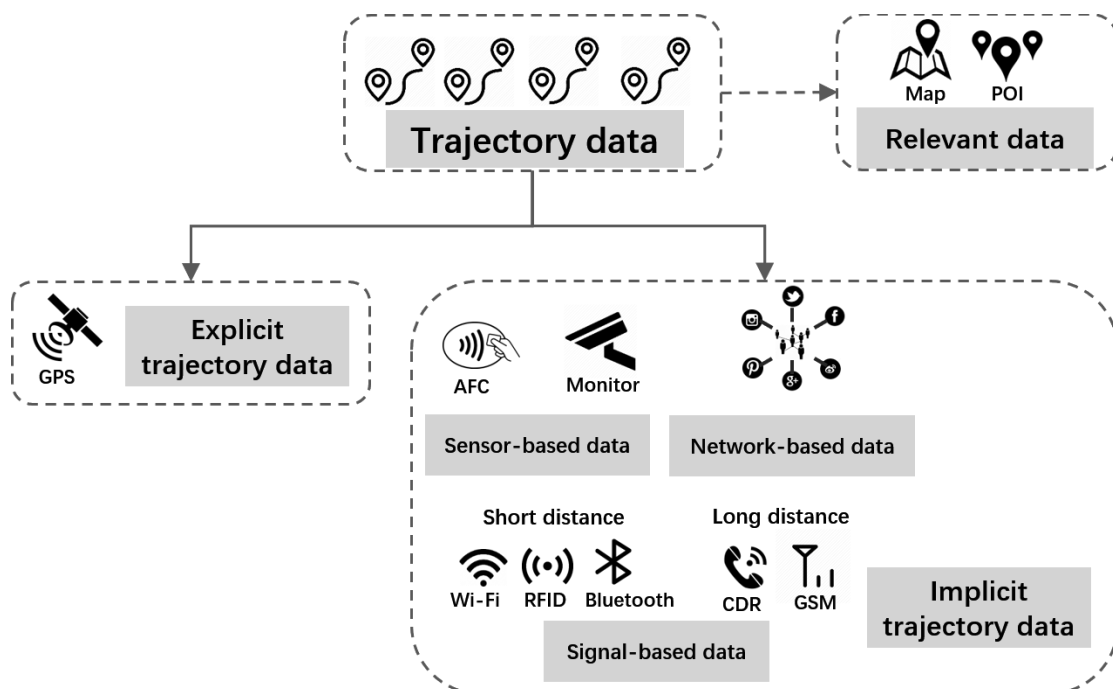
Υπάρχουν πολυάριθμες εφαρμογές και υπηρεσίες των δεδομένων κίνησης για κυβερνητικούς και εμπορικούς οργανισμούς αλλά και για άλλες χρήσεις π.χ. ακαδημαϊκή. Για να γίνει πιο κατανοητή η ανάλυση δεδομένων τροχιάς, αναφέρουμε επιγραμματικά τον ορισμό της τροχιάς, η οποία θεωρείται ως η διαδρομή που ακολουθεί ένα κινούμενο αντικείμενο μέσω του χώρου ως συνάρτηση του χρόνου. Η ανάλυση δεδομένων τροχιάς βοηθά στην μείωση του κόστους της διαχείρισης και στην εφαρμογή κατάλληλων στρατηγικών αστικού σχεδιασμού π.χ. παρακολούθηση ασταθών κινήσεων στα οχήματα όπως υπερβολική ταχύτητα, και όπισθεν κίνηση σε μονόδρομους. Ακόμη και εγκληματικές συμπεριφορές μπορούν να αναλυθούν μέσα από την ανάλυση δεδομένων κίνησης όπως π.χ. το μοτίβο εγκληματικών λειτουργιών σε ληστείες. Στις υπηρεσίες αυτές θα πρέπει να προστεθούν και τα μέσα μαζικής μεταφοράς όπως η αποδοτικότερη μετακίνηση των ταξί και των λεωφορείων που είναι ζωτικής σημασίας για τις μετακινήσεις των πολιτών.

Σύμφωνα με τον Zheng, (2015) ως δεδομένα κίνησης ορίζεται το πλήθος των παρατηρήσεων που δημιουργείται από την καταγραφή του ίχνους ενός κινούμενου

¹ Internet of things (IoT)

αντικειμένου με την βοήθεια χρονολογικά καταγεγραμμένων ή ταξινομημένων σημείων θέσης. Οι Mazimpraka και Timpf (2016) ορίζουν τα δεδομένα κίνησης παρόμοια με το Zheng αλλά με συγκεκριμένο γεωγραφικό προσδιορισμό και όχι με τον ορισμό αυθαίρετων σημείων κίνησης. Σε μια ποικιλία δεδομένων κίνησης, ορισμένοι τύποι δεδομένων μπορεί να αναφέρονται σε κινήσεις ατόμων, ζώων, αντικειμένων και ούτω καθεξής, όπως δεδομένα που προέρχονται από ένα GPS. Αν και δεν ανήκουν στη κοινώς αντιληπτή κατηγορία δεδομένων κίνησης η εξαγωγή κατάλληλων δεδομένων προς ανάλυση μπορεί να γίνει με συνδυαστικές μεθόδους όπως εφαρμογή μέτρων γεωγραφικής θέσης σε κοινωνικά ή βιομηχανικά κ.α. δεδομένα.

Επομένως, τα δεδομένα κίνησης μπορούν να κατηγοριοποιηθούν κατά προσέγγιση σε σαφή (explicit) δεδομένα τροχιάς και έμμεσα (implicit) δεδομένα κίνησης όπως φαίνεται και στην εικόνα 1. Ως βασικός εκπρόσωπος των άμεσων δεδομένων κίνησης μπορούν να θεωρηθούν τα δεδομένα που προέρχονται από την καταγραφή μέσω GPS σε σταθερά και ισαπέχοντα χρονικά διαστήματα. Αντίθετα, στα έμμεσα δεδομένα κίνησης, η μέτρηση των χρονικών διαστημάτων και των θέσεων είναι ασταθής και πιθανότατα τυχαία (Lee & Krumm, 2011).



Σχήμα 1 - Κατηγορίες δεδομένων κίνησης (Πηγή: : Lee & Krumm, 2011)

Σαφή δεδομένα (Explicit data)

Ορίζουμε ως σαφή δεδομένα κίνησης την συλλογή καλά δομημένων δεδομένων που παρέχουν άμεσα πληροφορίες χρόνου και τοποθεσίας και έχουν αρκετά ισχυρή χωρική και χρονική συνέχεια. Τα σαφή δεδομένα κίνησης βασίζονται στον χρόνο και τις τοποθεσίες και τα δεδομένα GPS είναι τα πιο αντιπροσωπευτικά δεδομένα, τα οποία χρησιμοποιούνται ευρέως σε έρευνες ανάλυσης της κίνησης. Κάθε εγγραφή δεδομένων GPS περιέχει πληροφορίες χρόνου, πλάτους και μήκους και άλλες πρόσθετες πληροφορίες, όπως η ταχύτητα (Kong et al, 2018).

Έμμεσα δεδομένα (Implicit data)

Σε αντίθεση με τα σαφή δεδομένα κίνησης, τα έμμεσα δεδομένα τροχιάς έχουν την αδύναμη χωρική και χρονική συνέχεια. Επιπλέον, τα ρητά δεδομένα τροχιάς συνήθως συγκαταλέγουν αρχεία δεδομένων με διαφορετικούς τύπους (format) αποθήκευσης.

Επιπλέον τα έμμεσα δεδομένα κίνησης ταξινομούνται με βάση τον τρόπο συλλογής τους σε βασιζόμενα σε σήματα (signal based), σε βασιζόμενα σε αισθητήρες (sensor based) και σε σήματα δικτύου (network based).

Δεδομένα σήματος

Η συλλογή δεδομένων βάσει σήματος απαιτεί πολλαπλές πηγές σήματος να διανέμονται σε διαφορετικές τοποθεσίες, π.χ. δρομολογητές για Wi-Fi, σταθμούς βάσης για εξοπλισμό GSM και ΔτΑ για Bluetooth και RFID. Απαιτούνται επίσης συσκευές λήψης, όπως κινητά τηλέφωνα. Μια μοναδική πηγή σήματος αντιστοιχεί σε πολλαπλές συσκευές λήψης. Επιπλέον, για δεδομένα βάσει σήματος, οι πηγές της απόκτησης δεδομένων είναι πολλές, ενώ η ακρίβεια τοποθεσίας είναι σχετικά χαμηλή. Τα πεδία δεδομένων βάσει σήματος τείνουν να είναι περίπλοκα, όπως η αναγνώριση συσκευής, χρόνος σύνδεσης / αποσύνδεσης, ισχύς σήματος και πρόσθετες πληροφορίες. Με βάση την απόσταση μετάδοσης σήματος, τα δεδομένα βάσει σήματος μπορούν να χωριστούν περαιτέρω σε δεδομένα μικρών αποστάσεων, συμπεριλαμβανομένου των δεδομένων Wi-Fi, δεδομένων Bluetooth και μεγάλων αποστάσεων όπως GSM δεδομένα.

Δεδομένα αισθητήρων

Για τα δεδομένα που βασίζονται στην καταγραφή από αισθητήρα, η καταγραφή αυτή γίνεται όταν το αντικείμενο περνά από τον αισθητήρα. Για παράδειγμα, η παρακολούθηση κυκλοφορίας μπορεί να συλλέξει πληροφορίες τοποθεσίας των οχημάτων που διέρχονται. Λόγω του φυσικού ορίου θέσης των αισθητήρων, το εύρος της απόκτησης δεδομένων είναι μικρό και η ακρίβεια της θέσης είναι σχετικά υψηλή. Τα δεδομένα που βασίζονται στον αισθητήρα καταγράφουν την ταυτότητα του αντικειμένου, το χρόνο μετάβασης και άλλες πρόσθετες πληροφορίες. Ως αντιπροσωπευτικά δεδομένα που βασίζονται σε αισθητήρες, είναι το TEO Pass που ανήκει στην οικογένεια των (Automatic Fare Collection-AFC) που συλλέγει δεδομένα συναλλαγών και μετακίνησης και χρησιμοποιείται για την ανάλυση ή τη βελτίωση των δημόσιων συγκοινωνιών (Pelletier et al., 2011).

Δεδομένα δικτύου

Με την άνοδο της χρήσης των κοινωνικών δικτύων τα τελευταία χρόνια, μπορούν να ληφθούν περισσότερα κοινωνικά δεδομένα με γεωγραφικές ετικέτες, τα οποία παρέχουν μια νέα αντίληψη για την ανάλυση δεδομένων κίνησης. Τέτοιου είδους δεδομένα μεταφορτώνονται από χρήστες και λαμβάνονται από ολόκληρο το Διαδίκτυο. Έτσι, το ορίζουμε ως δεδομένα που βασίζονται στο δίκτυο. Τα δεδομένα από το Facebook και το Weibo είναι αντιπροσωπευτικά αυτής της κατηγορίας δεδομένων που βασίζονται στο δίκτυο. Τα δεδομένα που βασίζονται στο δίκτυο είναι ιδιαίτερα ενδιαφέροντα και μπορούν να προσφέρουν πρόσθετες σημασιολογικές πληροφορίες για κοινωνικά συμβάντα, ωστόσο, λόγω της μεγάλης εξάρτησης από τις συμπεριφορές των χρηστών (μεροληψία), περιέχουν παρεμβολές (θόρυβο) και χρειάζονται πολύπλοκη προ επεξεργασία (φιλτράρισμα).

Άλλες μορφές δεδομένων κίνησης

Τα δεδομένα κίνησης περιέχουν ένα μεγάλο αριθμό πληροφοριών. Ωστόσο, στην ανάλυσή τους, οι ερευνητές συνήθως τα ενσωματώνουν με άλλα δεδομένα για να επιτύχουν αποτελεσματικότερη εξόρυξη δεδομένων κίνησης. Δύο κοινά σχετικά δεδομένα, δεδομένα είναι το Σημείο ενδιαφέροντος (Point of Interest-POI) και τα δεδομένα χάρτη (map data).

Το POI περιέχει πληροφορίες για πραγματικά κτίρια, όπως ξενοδοχεία, νοσοκομεία, super market και σταθμούς. Με την εφαρμογή των δεδομένων POI, μπορούμε να μετατρέψουμε τις πληροφορίες γεωγραφικού μήκους και πλάτους σε σημαντικές πληροφορίες δόμησης και να παρέχουμε μια λογική εξήγηση των προτύπων και των φαινομένων που ανακαλύφθηκαν στα δεδομένα κίνησης. Η περιοχή ενδιαφέροντος (Region of Interest - ROI), ο χώρος ενδιαφέροντος (Area of Interest-AOI) και ο όγκος ενδιαφέροντος (VOI) σχετίζονται ιδιαίτερα με τα δεδομένα POI.

Τα δεδομένα κίνησης περιέχουν μεγάλο αριθμό γεωγραφικών στοιχείων. Όμως τα αστικά οδικά δίκτυα και η κατανομή εδάφους έχουν τεράστιο αντίκτυπο στο σχηματισμό τροχιάς κινούμενων αντικειμένων όπως ατόμων και οχημάτων. Επομένως, είναι αναμενόμενο ως ιδέα και πρακτική να ενσωματώνουμε τα δεδομένα τροχιάς σε έναν χάρτη, ο οποίος περιέχει πλούσιες πληροφορίες, όπως οδικά δίκτυα για την προβολή της κατανομής αυτών των δεδομένων. Η σημασία των δεδομένων χάρτη για την ανάλυση δεδομένων κίνησης δεν είναι αμελητέα. Το Google Maps, το Leaflet API και το Mapbox API αποτελούν πρακτικές εφαρμογές αυτών.

1.3. Εφαρμογές των δεδομένων κίνησης

Σε σχέση με την ταξινόμηση των δεδομένων κίνησης, μπορούμε να διαπιστώσουμε ότι τα σαφή δεδομένα κίνησης, χρησιμοποιούνται ευρέως και συχνότερα στις εφαρμογές των δεδομένων κίνησης λόγω της σαφούς δομής και των πλούσιων χωρικών και χρονικών πληροφοριών. Χάρη ευκολίας συλλογής τους, η χρήση δεδομένων με βάση αισθητήρα και δεδομένων βάσει σήματος (μεγάλη απόσταση) είναι επίσης σχετικά υψηλή σε εφαρμογές. Ωστόσο, για δεδομένα βάσει σήματος (μικρή απόσταση), όπως δεδομένα Wi-Fi και δεδομένα Bluetooth, το χαρακτηριστικό του μικρού εύρους περιορίζει το εύρος των εφαρμογών του σε αναλύσεις μοτίβων ανθρώπινης κινητικότητας και χρονικών συμπερασμάτων. Σε σύγκριση με άλλες εφαρμογές, η ανίχνευση ανωμαλιών χρησιμοποιεί δεδομένα που βασίζονται σε αισθητήρες, ειδικά δεδομένα παρακολούθησης βίντεο, με υψηλή αξία εφαρμογής. Πρόβλεψη δραστηριοτήτων, εκτίμηση σκοπού ταξιδιού και αστικών λειτουργιών, εξαρτώνται ιδιαίτερα από τα δεδομένα τύπου POI. Και σχεδόν όλες οι εφαρμογές τροχιάς βασίζονται σε δεδομένα χάρτη.

Ως ένα από τα πιο σημαντικά ερευνητικά θέματα είναι η ανθρώπινη συμπεριφορά κίνησης η οποία έχει προσελκύσει την προσοχή πολλών μελετών λόγω της πολυπλοκότητας και της μεταβλητότητάς της για μεγάλο χρονικό διάστημα. Οι πληροφορίες που κρύβονται στα δεδομένα κίνησης μπορούν να βοηθήσουν στην κατανόηση της συμπεριφοράς του ανθρώπου και στη μεταφορά των ενδιαφερόντων μοτίβων κινητικότητας για τη βελτίωση της ποιότητας της αστικής ζωής.

Η ανάλυση του μοτίβου της ανθρώπινης κινητικότητας είναι ένας σημαντικός υποτομέας της ανάλυσης ανθρώπινης συμπεριφοράς και μπορεί να δώσει πληροφορίες για πολλαπλά ζητήματα όπως ο πολεοδομικός σχεδιασμός. Οι σύνδεσμοι μεταξύ των ατόμων είναι πολύ προβλέψιμοι, αλλά οι παράγοντες που επηρεάζουν κυρίως την ανθρώπινη κίνηση είναι ασαφείς. Οι Giannotti et al. (2013), συμφωνούν με την πολυπλοκότητα της ανθρώπινης κίνησης μέσα από την ανάλυση πολλαπλών μεταβλητών του μοτίβου αλλά και της πρόβλεψης της πιθανότητας ανθρώπινης κίνησης. Οι Shaw et al., (2016) συσχέτισαν την ανθρώπινη συμπεριφορά και τις σχέσεις με τις τροχιές στην εποχή των κινητών και στην ανάλυση των μεγάλων δεδομένων (big data).

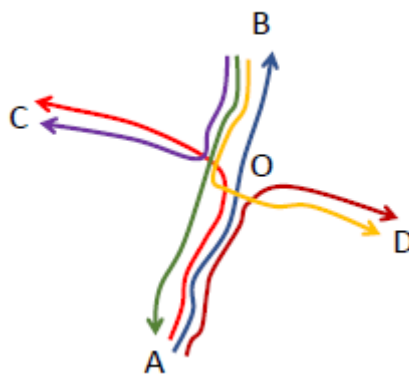
Με βάση συγκεκριμένους παράγοντες που επηρεάζουν την κινητικότητα του ανθρώπου και τον σαφή στόχο ανάλυσης, οι ερευνητές προτείνουν αποτελεσματικές μεθοδολογίες και πλαίσια για την κατανόηση του μοτίβου της ανθρώπινης κινητικότητας. Οι Renso et al. (2013), παρουσίασαν μια σημασιολογικά εμπλουτισμένη διαδικασία ανακάλυψης γνώσης για την εκτίμηση και την ουσιώδη ερμηνεία της ανθρώπινης κινητικότητας, με εφαρμογές στη διαχείριση της κυκλοφορίας και της ψυχαγωγίας. Οι Qiao et al. (2017), πρότειναν ένα αναλυτικό πλαίσιο της συμπεριφοράς της κίνησης για την ανάλυση της μαζικής κυκλοφορίας δεδομένων από το IoT όπως και την επικύρωση αυτού του πλαισίου με συγκεκριμένους κοινούς κανόνες.

Η ανθρώπινη κίνηση αποτελεί μια ευρεία έννοια και περιέχει βραχυπρόθεσμες, μεσοπρόθεσμες και μακροπρόθεσμες συμπεριφορές. Οι βραχυπρόθεσμες και μεσοπρόθεσμες συμπεριφορές έχουν υψηλή πιθανότητα και μεταβλητότητα, γεγονός που τις καθιστά πιο δύσκολες στην μελέτη σε σύγκριση με τις μακροπρόθεσμες

συμπεριφορές κίνησης. Ένα παράδειγμα μακροχρόνιας μελέτης είναι η μετακίνηση εργαζομένων προς την εργασία τους. Οι Zhou, Murphy και Long (2014) διεξήγαγαν μια λεπτομερή έρευνα σχετικά με τον τρόπο αύξησης της αποτελεσματικότητας των μετακινήσεων χρησιμοποιώντας δεδομένα έξυπνων καρτών (IC και ICC) στο Πεκίνο και υπολόγισαν το πλαίσιο της πλεονάζουσας μετακίνησης. Οι Ma et al. (2017), ανέλυσαν περαιτέρω το μοτίβο μετακίνησης των ατόμων λαμβάνοντας υπόψη πολλές μεταβλητές όπως ώρα αναχώρησης, απόσταση ταξιδιού και τον αριθμό των ημερών ταξιδιού. Επιπλέον, παρείχαν και μια προσέγγιση για τον εντοπισμό μετακινούμενων και διαμετακομιστικών μετακινήσεων.

1.4. Αντληση δεδομένων κίνησης

Μια τροχιά είναι η διαδρομή που ακολουθεί ένα κινούμενο αντικείμενο μέσω του χώρου ως συνάρτηση του χρόνου. Έτσι, μπορεί να γίνει αντιληπτή και ως μια χρονικά σφραγισμένη σειρά σημείων τοποθεσίας, που υποδηλώνεται ως $\{x_1, y_1, t_1, x_2, y_2, t_2, \dots, x_N, y_N, t_N\}$ όπου x_i, y_i αντιπροσωπεύουν γεωγραφικές συντεταγμένες του κινούμενου αντικειμένου τη στιγμή t_i και N είναι ο συνολικός αριθμός των στοιχείων στην σειρά. Για τη δημιουργία της τροχιάς, ένα κινούμενο αντικείμενο πρέπει να αποκτήσει τις συντεταγμένες του x, y τη στιγμή t (Lee & Krumm, 2011). Ένα παράδειγμα τροχιών παρουσιάζεται στην εικόνα 2. Στην συγκεκριμένη εικόνα παρουσιάζονται 6 τροχιές στους άξονες XY όπου κάθε μια αποτυπώνεται με ένα διαφορετικό ζεύγος εκκίνησης και προορισμού. Πιο συγκεκριμένα, τα παρατηρούμενα ζεύγη είναι $A \rightarrow B, A \rightarrow C, A \rightarrow D, B \rightarrow A, B \rightarrow C$ και $B \rightarrow D$. Αυτές οι 6 τροχιές έχουν τον ίδιο χρόνο εκκίνησης και παρόμοια ταχύτητα κίνησης (Tampakis et al., 2019)



Σχήμα 2 - Αποτύπωση 6 διαφορετικών τροχιών στους άξονες XY (Πηγή: Tampakis, et al., 2019)

Υπάρχουν πολλές τεχνολογίες εντοπισμού θέσης, π.χ. παγκόσμιο σύστημα εντοπισμού θέσης (GPS), το οποίο μπορεί να χρησιμοποιηθεί για τον προσδιορισμό της θέσης ενός κινούμενου αντικειμένου. Για παράδειγμα, τα περισσότερα από τα έξυπνα τηλέφωνα διαθέτουν ήδη ενσωματωμένο δέκτη GPS και έτσι μπορούν εύκολα να αντλήσουν τις δικές τους τοποθεσίες. Υπάρχουν περιπτώσεις που το GPS έχει χρησιμοποιήσει και σημεία πρόσβασης Wi-Fi ως το υποκείμενο σύστημα αναφοράς για τον προσδιορισμό της θέσης ενός κινούμενου αντικειμένου.

Επιπλέον, ακόμη και σε καταστάσεις όπου το σήμα GPS είναι κακό και δεν υπάρχει κοντινό σημείο πρόσβασης Wi-Fi για να χρησιμοποιηθεί ως εντοπισμός θέσης υπάρχουν οι τεχνικές «νεκρού υπολογισμού» (dead reckoning) που μπορούν να χρησιμοποιηθούν για την εκτίμηση της θέσης του κινούμενου αντικειμένου προχωρώντας από μια γνωστή θέση χρησιμοποιώντας τα στοιχεία πορεία, ταχύτητα, χρόνο και απόσταση που πρέπει να διανύσει. Ενώ οι παραδοσιακές μέθοδοι πλοήγησης του νεκρού υπολογισμού για την απόκτηση τοποθεσίας έχουν αντικατασταθεί από σύγχρονες τεχνολογίες εντοπισμού θέσης, εξακολουθεί να είναι μια πολύ χρήσιμη μέθοδος για τη συλλογή δεδομένων κίνησης, ειδικά όταν χαθεί η λήψη GPS π.χ. σε μια σήραγγα και γενικά όταν οι πιο μοντέρνες μέθοδοι δεν είναι εφαρμόσιμες.

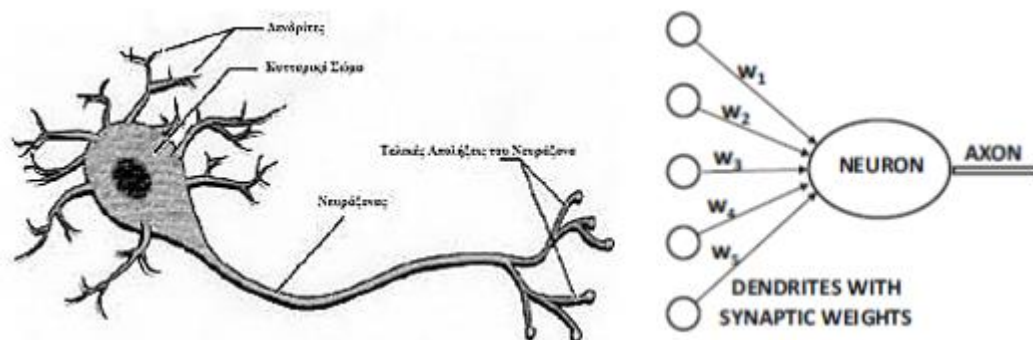
Δεδομένου ότι οι γεωγραφικές συντεταγμένες με χρονική σήμανση μπορούν να ληφθούν αυθαίρετα από ένα κινούμενο αντικείμενο, είναι σημαντικό να διευθετηθεί πριν από την συλλογή των δεδομένων, αν το κινούμενο αντικείμενο πρέπει να αναφέρει όλα τα δεδομένα τροχιάς του δείγματος στο διακομιστή τοποθεσίας για μεταφόρτωση στη βάση δεδομένων. Προφανώς η απάντηση εξαρτάται από τις ανάγκες και τους στόχους της έρευνας.

Δεδομένου ότι η απόκτηση δεδομένων πραγματοποιείται σε κινούμενο αντικείμενο, υποθέτουμε ότι η τοποθεσία που υποδεικνύει είναι η βέλτιστη και ότι τα παρεχόμενα δεδομένα έχουν την υψηλότερη ακρίβεια. Στην πραγματικότητα, η συλλογή πάντοτε περιέχει κάποια ανακρίβεια με βάση την μέθοδο συλλογής τους. Έτσι, η ακρίβεια των δεδομένων τοποθεσίας αναμένεται να είναι τόσο υψηλή όσο και τα όργανα συλλογής τους.

Κεφάλαιο 2. Νευρωνικά δίκτυα

2.1. Εισαγωγή

Τα τεχνητά νευρωνικά δίκτυα είναι μια δημοφιλής μέθοδος μηχανικής μάθησης που προσομοιάζει του μηχανισμούς μάθησης βιολογικών οργανισμών (Aggarwal, 2018). Το ανθρώπινο νευρικό σύστημα περιέχει κύτταρα τα οποία αναφέρονται ως νευρώνες. Οι νευρώνες συνδέονται μεταξύ τους με τη χρήση (νευρο)αξόνων (axons) και δενδριτών (dendrites) και οι περιοχές σύνδεσης μεταξύ τους αναφέρονται ως συνάψεις. Οι συνδέσεις αυτές απεικονίζονται στην εικόνα 3. Τα δυνατά σημεία των συναπτικών συνδέσεων συχνά μεταβάλλουν την αντίδραση σε εξωτερικούς ερεθισμούς. Αυτές οι μεταβολές αποτελούν και τον τρόπο μάθησης σε ζωντανούς οργανισμούς.



Σχήμα 3 - Βιολογικό και Τεχνητό Νευρωνικό Δίκτυο

Αυτός ο βιολογικός μηχανισμός προσομοιώνεται σε τεχνητά νευρωνικά δίκτυα, τα οποία περιέχουν υπολογίσιμες μονάδες που αναφέρονται ως νευρώνες. Οι υπολογιστικές μονάδες συνδέονται μεταξύ τους μέσω βαρών (weights ή w), τα οποία εξυπηρετούν τον ίδιο ρόλο με αυτό των συναπτικών συνδέσεων σε βιολογικούς οργανισμούς. Κάθε είσοδος σε έναν νευρώνα κλιμακώνεται με το βάρος, το οποίο επηρεάζει τη λειτουργία (function) που υπολογίζεται σε αυτήν τη μονάδα. Αυτή η αρχιτεκτονική απεικονίζεται στο Σχήμα 2.1 (B). Ένα τεχνητό νευρικό δίκτυο υπολογίζει μια συνάρτηση των εισόδων (inputs) πολλαπλασιάζοντας τις υπολογισμένες τιμές από τους νευρώνες εισόδου στους νευρώνες εξόδου (outputs) και χρησιμοποιώντας τα βάρη ως ενδιάμεσες παραμέτρους. Η μάθηση συμβαίνει αλλάζοντας τα βάρη που συνδέονται οι νευρώνες. Ακριβώς όπως χρειάζονται εξωτερικά ερεθίσματα για την μάθηση σε βιολογικούς οργανισμούς. Το εξωτερικό ερέθισμα στα τεχνητά νευρωνικά δίκτυα παρέχεται από τα δεδομένα εκπαίδευσης

(training data) που περιέχουν παραδείγματα ζευγών εισόδου-εξόδου της συνάρτησης προς μάθηση. Για παράδειγμα, το εκπαιδευτικό σετ δεδομένων (training set) ενδέχεται να περιέχει αναπαραστάσεις εικόνων (input) και τις επισημασμένες ετικέτες τους (π.χ. καρότο, μπανάνα) ως έξοδος (output). Αυτά τα ζεύγη αρχικών δεδομένων (training data) τροφοδοτούνται στο νευρωνικό δίκτυο χρησιμοποιώντας τις παραστάσεις εισόδου για να κάνουν προβλέψεις σχετικά με τις ετικέτες εξόδου. Τα δεδομένα παρέχουν ανατροφοδότηση σχετικά με την ορθότητα των βαρών στο νευρωνικό δίκτυο ανάλογα με το πόσο ακριβής ήταν η πρόβλεψη π.χ. πιθανότητα καρότου όταν η εικόνα συμβολίζει καρότο.

2.2. Ιστορική αναδρομή

Τα τελευταία 10 χρόνια, τα συστήματα τεχνητής νοημοσύνης με την καλύτερη απόδοση - όπως τα αναγνωριστικά ομιλίας σε smart phone ή ο τελευταίος αυτόματος μεταφραστής της Google - προέκυψαν από αυτές τις τεχνικές βαθιάς γνώσης.

Η βαθιά μάθηση είναι στην πραγματικότητα ένα νέο όνομα για μια προσέγγιση της τεχνητής νοημοσύνης που ονομάζεται νευρωνικά δίκτυα, τα οποία εισέρχονται και εξέρχονται από τη μόδα για περισσότερα από 70 χρόνια. Τα νευρωνικά δίκτυα προτάθηκαν για πρώτη φορά το 1944 από τους Warren McCullough και Walter Pitts, δύο ερευνητές του Πανεπιστημίου του Σικάγο που μετακόμισαν στο MIT το 1952 ως ιδρυτικά μέλη αυτού που μερικές φορές ονομάζεται το πρώτο γνωστικό τμήμα επιστημών (Lyon & Hafner, 1998).

Τα νευρωνικά δίκτυα ήταν ένας σημαντικός τομέας έρευνας τόσο στη νευροεπιστήμη όσο και στην επιστήμη των υπολογιστών μέχρι το 1969, όταν, σύμφωνα με τη γνώση της επιστήμης των υπολογιστών, αφαιρέθηκαν από τους μαθηματικούς του MIT Marvin Minsky και Seymour Papert, (1969) οι οποίοι ένα χρόνο αργότερα θα γίνουν συν-διευθυντές του νέο εργαστήριο Τεχνητής Νοημοσύνης του MIT.

Η τεχνική αυτή κίνησε ξανά το ενδιαφέρον των επιστημόνων κατά την δεκαετία του 1980, και επέστρεψε ξανά μετά το 2000, μετά την εμφάνιση επεξεργαστών μεγάλης ισχύος. Τα νευρωνικά δίκτυα είναι ένα μέσο για τη μηχανική εκμάθηση, στην οποία ένας υπολογιστής μαθαίνει να εκτελεί κάποια εργασία αναλύοντας παραδείγματα

εκπαίδευσης. Συνήθως, τα παραδείγματα έχουν επισημανθεί εκ των προτέρων. Ένα σύστημα αναγνώρισης αντικειμένων, για παράδειγμα, μπορεί να τροφοδοτείται με χιλιάδες ετικέτες από εικόνες αυτοκινήτων, σπιτιών, φλιτζανιών καφέ και ούτω καθεξής, και θα βρει οπτικά μοτίβα στις εικόνες που συσχετίζονται σταθερά με συγκεκριμένες ετικέτες.

Στον ανθρώπινο εγκέφαλο, ένα νευρωνικό δίκτυο αποτελείται από χιλιάδες ή και εκατομμύρια απλούς κόμβους επεξεργασίας που είναι πυκνά διασυνδεδεμένοι. Τα περισσότερα από τα σημερινά νευρωνικά δίκτυα είναι οργανωμένα σε στρώματα κόμβων και «τροφοδοτούνται προς τα εμπρός», που σημαίνει ότι τα δεδομένα μετακινούνται μέσω αυτών σε μία μόνο κατεύθυνση. Ένας μεμονωμένος κόμβος μπορεί να είναι συνδεδεμένος με πολλούς κόμβους στο επίπεδο κάτω από αυτό, από το οποίο λαμβάνει δεδομένα και αρκετούς κόμβους στο επίπεδο πάνω από αυτό, στο οποίο στέλνει δεδομένα.

Σε καθεμία από τις εισερχόμενες συνδέσεις του, ένας κόμβος θα εκχωρήσει έναν αριθμό γνωστό ως "βάρος". Όταν το δίκτυο είναι ενεργό, ο κόμβος λαμβάνει ένα διαφορετικό στοιχείο δεδομένων - έναν διαφορετικό αριθμό - σε κάθε μία από τις συνδέσεις του και τον πολλαπλασιάζει με το σχετικό βάρος. Στη συνέχεια προσθέτει τα εξαγόμενα γινόμενα μαζί, αποδίδοντας έναν μόνο αριθμό. Εάν αυτός ο αριθμός είναι κάτω από μια οριακή τιμή (threshold), ο κόμβος δεν μεταφέρει δεδομένα στο επόμενο επίπεδο. Εάν ο αριθμός υπερβαίνει την οριακή τιμή, τότε ο κόμβος «τροφοδοτεί», ο οποίος στα σημερινά νευρωνικά δίκτυα γενικά σημαίνει αποστολή του αριθμού - του αθροίσματος των σταθμισμένων εισόδων - σε όλες τις εξερχόμενες συνδέσεις του.

Όταν ένα νευρωνικό δίκτυο εκπαιδεύεται, όλα τα βάρη και οι οριακές του τιμές αρχικά ρυθμίζονται σε τυχαίες τιμές. Τα δεδομένα μάθησης τροφοδοτούνται στο κάτω στρώμα - το επίπεδο εισόδου - και περνούν μέσα από τα επόμενα στρώματα, πολλαπλασιάζονται και προστίθενται μαζί με πολύπλοκους τρόπους, έως ότου φτάσει τελικά, ριζικά μεταμορφωμένο, στο επίπεδο εξόδου. Κατά τη διάρκεια της εκπαίδευσης (training), τα βάρη και τα κατώτατα όρια προσαρμόζονται συνεχώς έως ότου τα δεδομένα εκπαίδευσης με τις ίδιες ετικέτες αποδίδουν σταθερά παρόμοιες εξόδους.

Τα νευρωνικά δίκτυα που περιγράφονται από τους McCulloch και Pitts, (1944) είχαν κατώφλια και βάρη, αλλά δεν ήταν διατεταγμένα σε στρώματα και οι ερευνητές δεν διευκρίνισαν κανένα μηχανισμό εκπαίδευσης. Αυτό που έδειξαν οι McCulloch και Pitts ήταν ότι ένα νευρωνικό δίκτυο θα μπορούσε, καταρχήν, να υπολογίσει οποιαδήποτε λειτουργία παρόμοια με ένα αυτόματο υπολογιστικό σύστημα. Το αποτέλεσμα τους αναφερόταν περισσότερο στην νευροεπιστήμη παρά στην επιστήμη των υπολογιστών, δείχνοντας όμως ότι ο ανθρώπινος εγκέφαλος θα μπορούσε να θεωρηθεί ως μια υπολογιστική συσκευή.

Τα νευρωνικά δίκτυα συνεχίζουν να αποτελούν πολύτιμο εργαλείο για τη νευροεπιστημονική έρευνα. Για παράδειγμα, συγκεκριμένες διατάξεις δικτύου ή κανόνες για την προσαρμογή των βαρών και των ορίων έχουν αναπαράγει παρατηρούμενα χαρακτηριστικά της ανθρώπινης νευροανατομίας και της γνώσης, μια ένδειξη ότι καταγράφουν κάτι για το πώς ο εγκέφαλος επεξεργάζεται πληροφορίες.

Το πρώτο εκπαιδευμένο νευρωνικό δίκτυο, το Perceptron, επιδείχθηκε από τον ψυχολόγο του Πανεπιστημίου Cornell, Frank Rosenblatt, το 1957. Ο σχεδιασμός του Perceptron ήταν παρόμοιος με αυτόν ενός σύγχρονου νευρωνικού δικτύου, εκτός από το ότι είχε μόνο ένα στρώμα με ρυθμιζόμενα βάρη και κατώτατα όρια, ανάμεσα σε είσοδο και επίπεδα εξόδου (Hardesty, 2017).

Οι Perceptrons ήταν ένας ενεργός τομέας έρευνας τόσο στην ψυχολογία όσο και στην επιστήμη των υπολογιστών μέχρι το 1959, όταν ο Minsky και ο Papert δημοσίευσαν ένα βιβλίο με τίτλο "Perceptrons", το οποίο απέδειξε ότι η εκτέλεση ορισμένων αρκετά κοινών υπολογισμών στα Perceptrons θα ήταν πρακτικά χρονοβόρα (Minsky & Papert, 1969).

Μέχρι τη δεκαετία του 1980, ωστόσο, οι ερευνητές είχαν αναπτύξει αλγόριθμους για την τροποποίηση των βαρών και των κατωφλίων νευρωνικών δικτύων που ήταν αρκετά αποτελεσματικοί για δίκτυα με περισσότερα από ένα επίπεδα, αφαιρώντας πολλούς από τους περιορισμούς που εντοπίστηκαν από τους Minsky και Papert.

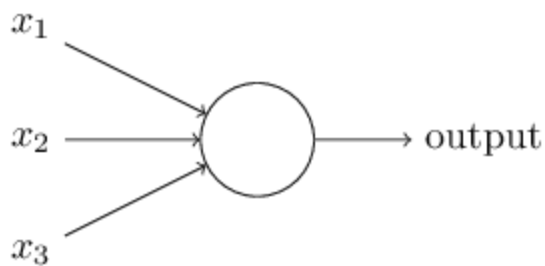
Τα τελευταία χρόνια, οι ερευνητές στην επιστήμη των υπολογιστών έχουν αρχίσει να επινοούν έξυπνες μεθόδους για την εξαγωγή των αναλυτικών στρατηγικών που υιοθετούν τα νευρωνικά δίκτυα. Αλλά στη δεκαετία του 1980, οι στρατηγικές των δικτύων ήταν ανεξάρτητες. Έτσι, γύρω στα τέλη του αιώνα, τα νευρωνικά δίκτυα αντικαταστάθηκαν από τεχνικές Support Vector Machines, μια εναλλακτική προσέγγιση στη μηχανική μάθηση που βασίζεται σε ορισμένα πολύ ανώτερα.

Η πρόσφατη αναζωπύρωση στα νευρωνικά δίκτυα - η επανάσταση της βαθιάς μάθησης - προέρχεται από τη βιομηχανία παιχνιδιών υπολογιστών. Οι περίπλοκες εικόνες και ο γρήγορος ρυθμός των σημερινών βιντεοπαιχνιδιών απαιτούν υλικό που μπορεί να διατηρηθεί και το αποτέλεσμα ήταν η μονάδα επεξεργασίας γραφικών (GPU), η οποία συσκευάζει χιλιάδες σχετικά απλούς πυρήνες επεξεργασίας σε ένα μόνο τσιπ. Δεν χρειάστηκε πολύς χρόνος για τους ερευνητές να συνειδητοποιήσουν ότι η αρχιτεκτονική μιας GPU είναι παρόμοια με αυτή ενός νευρωνικού δικτύου (Hardesty, 2017).

Οι σύγχρονες GPU επέτρεψαν στα δίκτυα ενός επιπέδου της δεκαετίας του 1960 και τα δίκτυα δύο έως τριών επιπέδων της δεκαετίας του 1980 να εξελιχθούν στα δίκτυα 10-, 15-, ακόμη και των 50 επιπέδων του σήμερα. Σε αυτό αναφέρεται το «βαθύ» στη «βαθιά μάθηση – Deep learning» - το βάθος των επιπέδων του δικτύου. Και επί του παρόντος, η βαθιά γνώση είναι υπεύθυνη για τα συστήματα με την καλύτερη απόδοση σε σχεδόν όλους τους τομείς της έρευνας τεχνητής νοημοσύνης.

2.3. Παραδείγματα νευρωνικών δικτύων

Σε αυτή την παράγραφο παρουσιάζουμε μερικά απλά νευρωνικά δίκτυα. Ας υποθέσουμε ότι έχουμε ένα δίκτυο που παίρνει αρκετές δυαδικές εισόδους, X_1, X_2, \dots και παράγει μία μόνο δυαδική έξοδο. Στο παράδειγμα που παρουσιάζεται στην εικόνα 4 έχουμε τρεις εισόδους, X_1, X_2, X_3 . Σε γενικές γραμμές θα μπορούσε να έχει περισσότερες ή λιγότερες εισόδους.



Σχήμα 4 - Νευρωνικό δίκτυο

Πηγή: <http://neuralnetworksanddeeplearning.com/chap1.htm>.

Ο Rosenblatt, (1957) πρότεινε έναν απλό κανόνα για τον υπολογισμό του εξαγόμενου αποτελέσματος. Εισήγαγε βάρη, w_1, w_2, \dots , που είναι πραγματικοί αριθμοί και εκφράζουν τη σημασία των αντίστοιχων εισόδων στην έξοδο. Η έξοδος του νευρώνα, είτε 0 είτε 1, καθορίζεται από το αν το σταθμισμένο άθροισμα $\sum w_j x_j$ είναι μικρότερο ή μεγαλύτερο από κάποια οριακή τιμή (threshold value). Όπως και τα βάρη, έτσι και η οριακή τιμή είναι ένας πραγματικός αριθμός και είναι μια παράμετρος του νευρώνα. Για να το θέσουμε με πιο ακριβείς αλγεβρικούς όρους:

$$output = \begin{cases} 0 & \text{εάν } \sum_j w_j x_j \leq threshold \\ 1 & \text{εάν } \sum_j w_j x_j > threshold \end{cases}$$

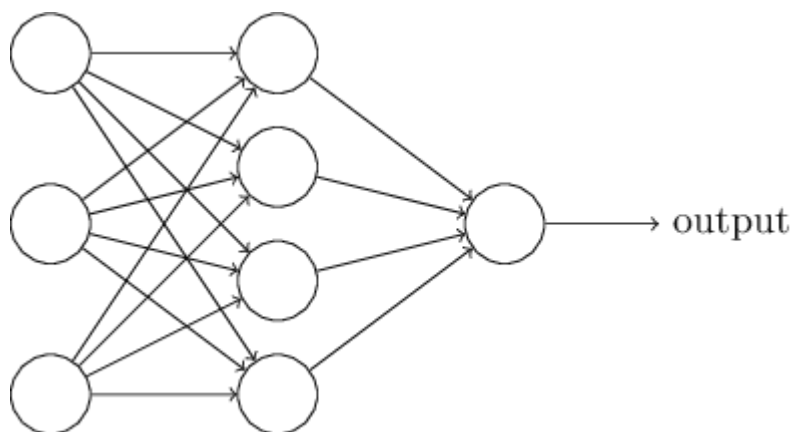
Αυτό είναι και το βασικό μαθηματικό μοντέλο. Ένας τρόπος με τον οποίο το μοντέλο αυτό μπορεί να γίνει αντιληπτό είναι ως μια συσκευή που λαμβάνει αποφάσεις σταθμίζοντας στοιχεία. Ένα ακόμη παράδειγμα είναι η λήψη απόφασης βάσει παραμέτρων που επηρεάζουν αυτή την απόφαση. Πιο συγκεκριμένα έστω ότι έχουμε την λήψη της απόφασης, μετάβαση ή μη σε μια συγκεκριμένη περιοχή με τις παραμέτρους

1. Καιρός
2. Συνοδεία και
3. Ευκολία μετακίνησης

Έστω ότι X_1, X_2 και X_3 οι μεταβλητές που αντιπροσωπεύουν αυτές τις τρεις παραμέτρους και το θετικό ενδεχόμενο παίρνει την τιμή 1 π.χ. $X_1=1$ σημαίνει ο καιρός θα είναι καλός. Μια πιθανή μοντελοποίηση αυτού είναι η $w_1=6$ και $w_2=2$ και $w_3=2$, δηλαδή ο καιρός είναι παράμετρος μεγαλύτερης βαρύτητας από τις υπόλοιπες 2. Εάν η οριακή τιμή οριστεί ως 5 τότε μόνο η εμφάνιση καλοκαιρίας ($x_1=1$) είναι αρκετή για την λήψη

της απόφασης. Μεταβολή της οριακής τιμής προς τα επάνω σημαίνει δυσκολία μετάβασης και προς τα κάτω ευκολία μετάβασης.

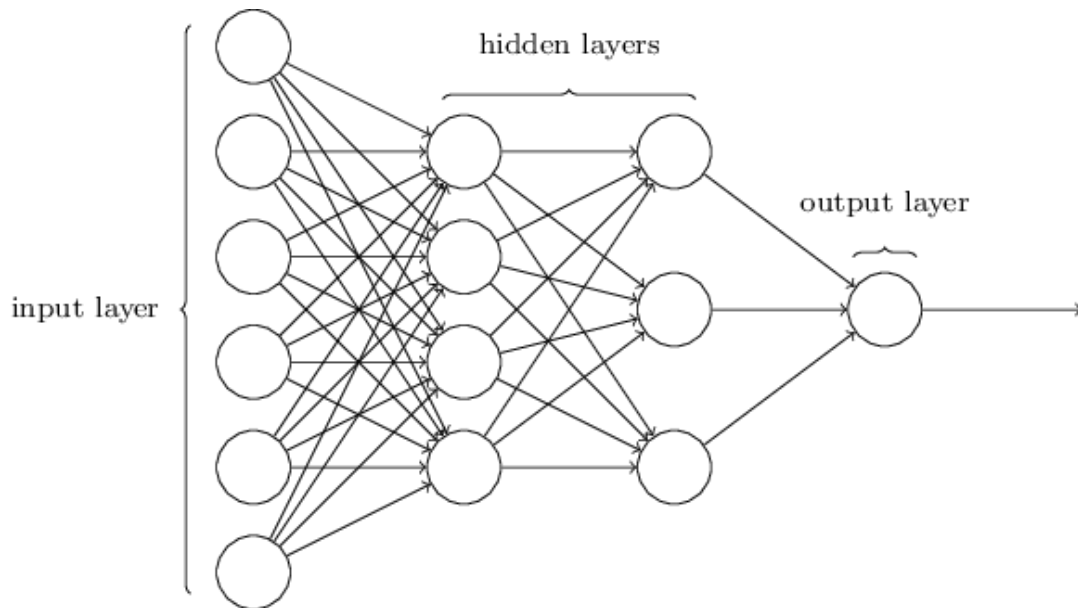
Ένα άλλο πιο σύνθετο μοντέλο παρουσιάζεται στην εικόνα 5. Το αριστερότερο στρώμα σε αυτό το δίκτυο ονομάζεται στρώμα εισόδου και οι νευρώνες εντός του στρώματος ονομάζονται νευρώνες εισόδου. Το δεξιότερο ή το επίπεδο εξόδου περιέχει τους νευρώνες εξόδου ή, όπως στην περίπτωση αυτή, έναν μόνο νευρώνα εξόδου. Το μεσαίο στρώμα ονομάζεται κρυφό στρώμα (hidden layer), δεδομένου ότι οι νευρώνες σε αυτό το στρώμα δεν είναι ούτε εισοδοί ούτε έξοδοι. Ο όρος «κρυμμένος» δεν σημαίνει τίποτα περισσότερο από το «όχι μια είσοδος ή μια έξοδος».



Σχήμα 5 - Νευρωνικό δίκτυο με ένα κρυφό επίπεδο

Πηγή: <http://neuralnetworksanddeeplearning.com/chap1.html>.

Το παραπάνω δίκτυο έχει μόνο ένα κρυφό επίπεδο, αλλά ορισμένα δίκτυα έχουν περισσότερα κρυφά επίπεδα. Για παράδειγμα, το ακόλουθο δίκτυο τεσσάρων επιπέδων έχει δύο κρυφά επίπεδα (εικόνα 6).



Σχήμα 6 - Νευρωνικό δίκτυο με δύο κρυφά επίπεδα και περισσότερες από μιας εξόδους

Πηγή: <http://neuralnetworksanddeeplearning.com/chap1.html>.

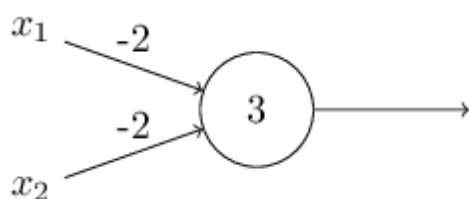
Σε αυτό το δίκτυο, η πρώτη στήλη του δικτύου λαμβάνει τρεις αποφάσεις, σταθμίζοντας τα στοιχεία εισόδου. Στο δεύτερο στρώμα κάθε κόμβος λαμβάνει μια απόφαση σταθμίζοντας τα αποτελέσματα από το πρώτο στρώμα λήψης αποφάσεων. Με αυτόν τον τρόπο, το δίκτυο στο δεύτερο στρώμα μπορεί να πάρει μια απόφαση σε πιο περίπλοκο και πιο αφηρημένο επίπεδο από ότι στο πρώτο στρώμα. Και ακόμη πιο περίπλοκες αποφάσεις μπορούν να ληφθούν από το δίκτυο στο τρίτο επίπεδο. Με αυτόν τον τρόπο, ένα δίκτυο πολλών επιπέδων μπορεί να βοηθήσει σε πιο εξελιγμένους υπολογισμούς για την λήψη αποφάσεων.

Και στις δύο προηγούμενες περιπτώσεις θεωρήσαμε μόνο μια πιθανή έκβαση (output). Παρόλα αυτά υπάρχουν περιπτώσεις που ζητούνται περισσότερα από ένα αποτελέσματα. Στην πραγματικότητα, η έξοδος παραμένει ενιαία. Τα πολλαπλά βέλη εξόδου είναι απλώς ένας χρήσιμος τρόπος ένδειξης ότι η έξοδος από ένα κόμβο χρησιμοποιείται ως είσοδος σε άλλους. Σε αυτή την περίπτωση η εξίσωση μπορεί να ξαναγραφτεί ως

$$output = \begin{cases} 0 & \text{εάν } \sum_j w_j x_j + b \leq 0 \\ 1 & \text{εάν } \sum_j w_j x_j + b > 0 \end{cases}$$

όπου πλέον με b συμβολίζουμε την αρνητική οριακή τιμή ανά περίπτωση που ονομάζεται και μεροληψία δικτύου (perceptron bias). Η τιμή αυτή είναι ένα μέτρο του πόσο εύκολο είναι να γίνει έναρξη της διαδικασίας απόφαση. Για ένα δίκτυο με πολύ μεγάλη τιμή b , είναι εξαιρετικά εύκολο το output να εξάγει τη τιμή 1. Αντίθετα εάν η τιμή είναι αρνητική και μεγάλη, τότε είναι δύσκολο το output να εξάγει την τιμή 1.

Κάπως συγκεκριμένα, και για ιστορικούς λόγους, τέτοια δίκτυα πολλαπλών επιπέδων ονομάζονται μερικές φορές πολυεπίπεδα perceptrons ή (Multilayer Perceptrons – MLPs), Εκτός όμως από την λήψη αποφάσεων τα νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν και ως κόμβοι λογικών τελεστών. Για παράδειγμα έστω ότι έχουμε τον απλό κόμβο της εικόνας 7 με βάρος-2 και $b=3$. Τότε η είσοδος 00 παράγει έξοδο 1 καθώς $(-2) * 0 + (-2) * 0 + 3 = 3$ είναι θετικό. Παρόμοιοι υπολογισμοί δείχνουν ότι οι εισοδοί 01 και 10 παράγουν έξοδο 1. Αλλά η είσοδος 11 παράγει έξοδο 0 καθώς $(-2) * 1 + (-2) * 1 + 3 = -1$ είναι αρνητικό. Έτσι το δίκτυο του συγκεκριμένου παραδείγματος εφαρμόζει μια NAND² πύλη.



Σχήμα 7 - Χρήση νευρωνικού δικτύου για τον υπολογισμό AND, OR και NAND

Πηγή: <http://neuralnetworksanddeeplearning.com/chap1.htm>.

Στο συγκεκριμένο κεφάλαιο κάναμε μια γενική περιγραφή της βαθιάς μάθησης και των νευρωνικών δικτύων. Σκοπός αυτού του κεφαλαίου ήταν η απαραίτητη εισαγωγή του αναγνώστη στις έννοιες του αντικειμένου πριν την χρήση πιο σύνθετων μεθόδων με εφαρμογή σε δεδομένα κίνησης που παρουσιάζονται στο επόμενο κεφάλαιο.

2.4. Μοντέρνες τεχνικές νευρωνικών δικτύων

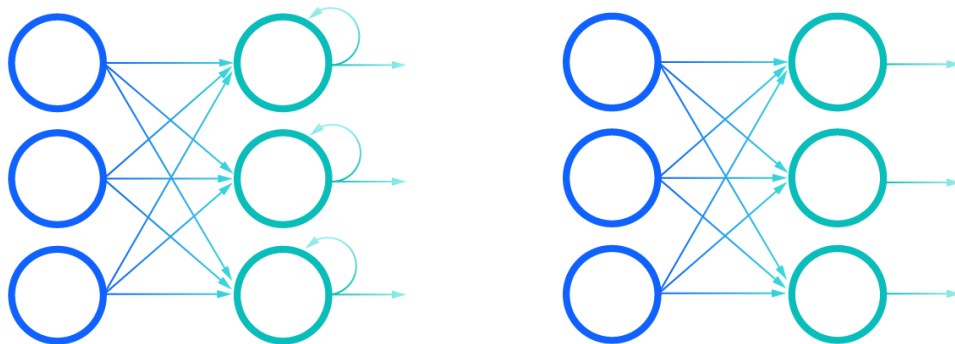
Σε αυτή την παράγραφο παρουσιάζονται οι πιο μοντέρνες τεχνικές νευρωνικών δικτύων Recurrent Neural Networks and LSTM.

² Not And: Εξάγει False Εάν και οι δύο τιμές εισόδου είναι True

2.4.1. Recurrent Neural Networks

Ένα επαναλαμβανόμενο νευρωνικό δίκτυο (RNN) είναι ένας τύπος τεχνητού νευρωνικού δικτύου που χρησιμοποιεί διαδοχικά δεδομένα ή δεδομένα χρονοσειρών. Αυτοί οι αλγόριθμοι βαθιάς μάθησης χρησιμοποιούνται συνήθως για τακτικά ή χρονικά προβλήματα, όπως μετάφραση γλώσσας, επεξεργασία φυσικής γλώσσας (nlp), αναγνώριση ομιλίας και ανάλυση δεδομένων κίνησης. Οι εφαρμογές τους είναι πάρα πολλές καθώς έχουν ενσωματωθεί σε δημοφιλείς εφαρμογές όπως το Siri, η φωνητική αναζήτηση και μετάφραση της Google.

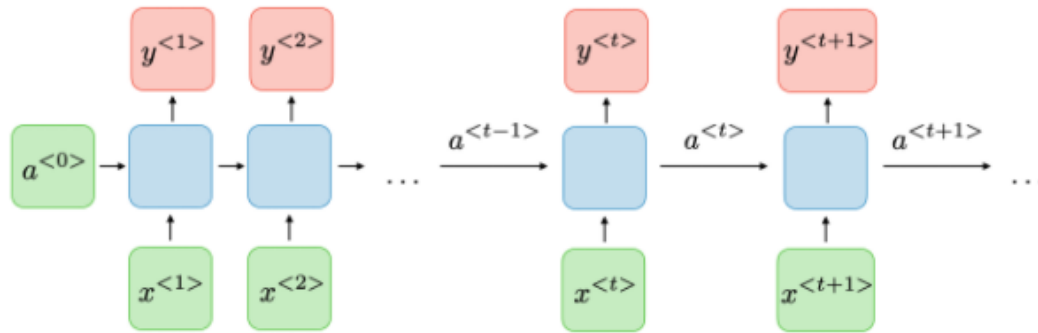
Όπως και άλλα νευρωνικά δίκτυα π.χ. Convolution Neural Networks ή CNN, τα επαναλαμβανόμενα νευρωνικά δίκτυα χρησιμοποιούν αρχικά δεδομένα (training data sets) για εκμάθηση. Χαρακτηριστικό τους γνώρισμα είναι η μνήμη τους καθώς λαμβάνουν πληροφορίες από προηγούμενες εισόδους για να επηρεάσουν την τρέχουσα είσοδο και έξοδο. Ενώ τα παραδοσιακά βαθιά νευρωνικά δίκτυα υποθέτουν ότι οι εισοδοί και οι εξοδοί είναι ανεξάρτητες μεταξύ τους, η έξοδος των επαναλαμβανόμενων νευρωνικών δικτύων εξαρτάται από τα προηγούμενα στοιχεία της ακολουθίας (IBM, 2020)



Σχήμα 8 - Αναπαράσταση των επαναλαμβανόμενων νευρωνικών δικτύων (A) και των νευρωνικών δικτύων Feedforward (B). (Πηγή: IBM, 2020)

Η μαθηματική αναπαράσταση αυτής της μεθόδου παρουσιάζεται στην εικόνα 9 όπου για κάθε χρονική στιγμή t η ενεργοποίηση $a^{<t>}$ και το εξαγόμενο $y^{<t>}$ εκφράζονται από τις σχέσεις

$$a^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_2) \text{ και } y^{<t>} = g_2(W_{ya}a^{<t>} + b_y)$$



Σχήμα 9 - Μαθηματική αναπαράσταση της μεθόδου RNN (Πηγή: Afshine & Shervine, 2020)

Τα πλεονεκτήματα αυτής της μεθόδου είναι ότι (Afshine & Shervine, 2020)

- Μπορούν να επεξεργαστούν δεδομένα ανεξαρτήτως του μεγέθους τους.
- Το μέγεθος του μοντέλου δεν αυξάνει από το μέγεθος των δεδομένων
- Οι υπολογισμοί λαμβάνουν υπόψη προηγούμενες πληροφορίες
- Τα βάρη διαμοιράζονται κατά μήκος της εξεταζόμενης χρονικής περιόδου

Αντίθετα, τα μειονεκτήματα αυτής της μεθόδου αναφέρονται ότι είναι (Afshine & Shervine, 2020)

- Χαμηλή ταχύτητα υπολογισμών
- Δυσκολία ανάκτησης της πληροφορίας στην βραχυπρόθεσμη μνήμη
- Δεν επιδέχεται επιπλέον δεδομένα κατά τους υπολογισμούς

2.4.2. LSTM

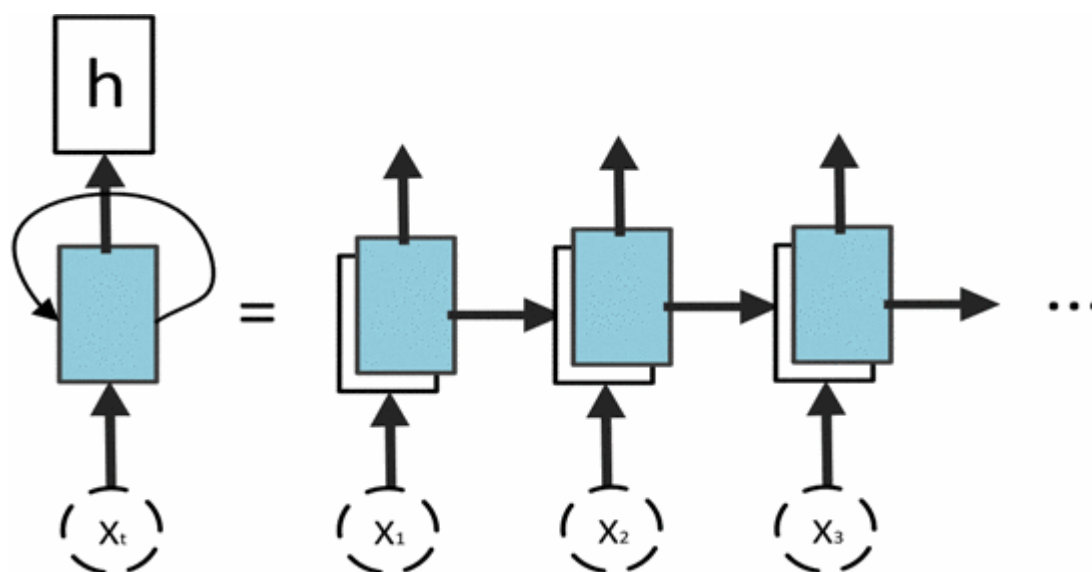
Όπως έχει ήδη αναφερθεί στην προηγούμενη παράγραφο, ένα μειονέκτημα της μεθόδου RNN είναι η δυσκολία ανάκτησης της πληροφορία στην βραχυπρόθεσμη μνήμη δηλαδή εάν μια ακολουθία είναι αρκετά μεγάλη, υπάρχει δυσκολία στην μεταφορά πληροφοριών από τα προηγούμενα βήματα στα μεταγενέστερα. Ένα παράδειγμα είναι η ελλιπής μεταφορά πληροφοριών σε text mining and predicting μιας παραγράφου όπου η μέθοδος RNN ενδέχεται να παραλείψει σημαντικές πληροφορίες από την αρχή της.

Για αυτό τον λόγο κατασκευάστηκε η μέθοδος LSTM (Long-Short Term Memory) η οποία αποτελείται από μια γραμμική διάταξη καταστάσεων που υποδέχονται τα inputs στην εξέλιξη του χρόνου t . Σε κάθε βήμα T μια σειρά πληροφοριών (input vectors)

εισάγεται στο LSTM και το εξαγόμενο αποτέλεσμα υπολογίζεται σύμφωνα με την σχέση (Balouji et al., 2018)

$$h_t = f_w(h_{t-1}, x_t)$$

όπου x_t είναι το input vector, h_t και h_{t-1} εκφράζουν τα διανύσματα της τρέχουσας κατάστασης τις χρονικές στιγμές t και $(t-1)$ και f_w εκφράζει μια μη γραμμική συνάρτηση ενεργοποίησης όπου το w αναφέρεται στις παραμέτρους των βαρών. Μια συνήθης γραφική απεικόνιση της μεθόδου LSTM παρουσιάζεται στην εικόνα 10.



Σχήμα 10 - Γραφική απεικόνιση της μεθόδου 2.7. (Πηγή: Balouji et al., 2018)

Τα μοντέλα LSTM πρέπει να εκπαιδευτούν με ένα σύνολο δεδομένων εκπαίδευσης πριν από την εφαρμογή τους σε πραγματικές εφαρμογές. Μερικές από τις πιο συνήθεις εφαρμογές της μεθόδου είναι οι εξής (GG, 2020) :

- Μοντελοποίηση γλωσσών ή δημιουργία κειμένου, που περιλαμβάνει τον υπολογισμό των λέξεων όπου μια ακολουθία λέξεων τροφοδοτείται ως είσοδος. Τα γλωσσικά μοντέλα μπορούν να χρησιμοποιηθούν σε επίπεδο χαρακτήρων, σε επίπεδο προτάσεων ή ακόμη και σε επίπεδο παραγράφου.
- Επεξεργασία εικόνας, που περιλαμβάνει την ανάλυση μιας εικόνας και την εξαγωγή λογικού συμπεράσματος σε μια πρόταση.
- Αναγνώριση ομιλίας και γραφής
- Μετάφραση γλώσσας

Επίσης ως μειονεκτήματα της μεθόδου έχει αναφερθεί ότι (GG, 2020) αν και επιλύουν το πρόβλημα της βραχυπρόθεσμης μνήμης εντούτοις δεν μπορούν να το αφαιρέσουν εντελώς. Έτσι είναι αναμενόμενες οι απώλειες δεδομένων αλλά σε αρκετά χαμηλότερο βαθμό σε σύγκριση με την μέθοδο RNN. Επίσης έχει παρατηρηθεί ότι η μέθοδος LSTM επηρεάζεται από την τυχαιότητα επιλογής του αρχικού βάρους. Μια συνήθης τεχνική για την επίλυση αυτού του προβλήματος είναι η επιλογή αρχικών βαρών με μικρό μέγεθος. Τέλος η μέθοδος LSTM είναι επιρρεπής σε υπερβολική τοποθέτηση παραμέτρων (overfitting).

Κεφάλαιο 3. Αλγόριθμοι Συσταδοποίησης

Στο κεφάλαιο αυτό θα αναφερθούν οι βασικοί αλγόριθμοι συσταδοποίησης που χρησιμοποιούνται στο κεφάλαιο 5, στην υλοποίηση του πρακτικού μέρους της διπλωματικής.

3.1 Ο αλγόριθμος k-Means

Σκοπός του k-means είναι η εύρεση των κέντρων των ομάδων σε ένα σετ δεδομένων όπου η απόσταση των κέντρων από τα σημεία του cluster που ανήκει το κέντρο, θέλουμε να είναι η ελάχιστη. Έχει 3 βασικά βήματα, ορισμός K κέντρων συστάδων με τυχαίο τρόπο, εισαγωγή αντικειμένου στη συστάδα με το πιο κοντινό κέντρο, ανανέωση του κέντρου της συστάδας. Τα δύο τελευταία βήματα επαναλαμβάνονται μέχρι να συγκλίνει ο αλγόριθμος. Η αντικειμενική συνάρτηση k-means είναι ένας από τους πιο δημοφιλείς στόχους ομαδοποίησης. Στην k-means τα δεδομένα χωρίζονται σε ασύνδετα σύνολα C_1, \dots, C_k όπου κάθε C_i αντιπροσωπεύεται από ένα κέντρο μάζας μ_i . Υποτίθεται ότι το σύνολο εισόδου \mathcal{X} είναι ενσωματωμένο σε κάποιο μεγαλύτερο μετρικό χώρο (\mathcal{X}', d) (έτσι ώστε $\mathcal{X} \subseteq \mathcal{X}'$) και τα κέντρα μάζας είναι μέλη του \mathcal{X}' . Η αντικειμενική συνάρτηση k-means μετρά την τετραγωνισμένη απόσταση μεταξύ κάθε σημείου στο \mathcal{X} έως το κέντρο μάζας του συμπλέγματος του. Το κέντρο μάζας του C_i ορίζεται ότι είναι

$$\mu_i(C_i) = \operatorname{argmin}_{\mu \in \mathcal{X}'} \sum_{x \in C_i} d(x, \mu)^2$$

Στη συνέχεια, ο στόχος του k-means είναι η ελαχιστοποίηση του σφάλματος J

$$J = \sum_{j=1}^k \sum_{i=1}^n (\|x_i^{(j)} - c_j\|)^2$$

, όπου k είναι ο αριθμός των ομάδων, n τα σημεία του cluster k, και τα x,y είναι το σημείο που ανήκει στο cluster k και το cluster k αντίστοιχα.

3.2 Φασματική Ομαδοποίηση

Συχνά, ένας βολικός τρόπος για την αναπαράσταση των σχέσεων μεταξύ σημείων σε ένα σύνολο δεδομένων $\mathcal{X} = \{x_1, \dots, x_m\}$, είναι με ένα γράφημα ομοιότητας όπου κάθε κορυφή αντιπροσωπεύει ένα σημείο δεδομένων x_i και κάθε δύο κορυφές συνδέονται με μια άκρη της οποίας το βάρος είναι η ομοιότητά τους, $W_{i,j} = s(x_i, x_j)$, όπου $W \in$

$\mathbb{R}^{m,m}$. Για παράδειγμα, μπορεί να οριστεί $W_{i,j} = \exp(-d(x_i, x_j)^2/\sigma^2)$ όπου $d(\cdot, \cdot)$ είναι μία συνάρτηση απόστασης και το σ είναι μία παράμετρος. Το πρόβλημα της ομαδοποίησης μπορεί τώρα να διατυπωθεί ως εξής: Έστω ότι ζητείται ένα διαμέρισμα του γραφήματος έτσι ώστε οι άκρες μεταξύ διαφορετικών ομάδων να έχουν χαμηλό βάρος και οι άκρες εντός μια ομάδας να έχουν υψηλά βάρη. Στους στόχους ομαδοποίησης που περιεγράφηκαν προηγουμένως, το επίκεντρο ήταν στη μία πλευρά του διαισθητικού ορισμού της ομαδοποίησης - διασφαλίζοντας με αυτό το τρόπο ότι τα σημεία στο ίδιο σύμπλεγμα είναι παρόμοια. Τώρα παρουσιάζονται στόχοι που επικεντρώνονται στην αντίθετη απαίτηση – δηλαδή τα σημεία που χωρίζονται σε διαφορετικά συμπλέγματα, να είναι ανόμοια μεταξύ τους.

Ο αλγόριθμος φασματικής ομαδοποίησης (Shalev-Shwartz & Ben-David, 2014) ξεκινά με την εύρεση του πίνακα H των ιδιοδιανυσμάτων k , που αντιστοιχούν στις μικρότερες ιδιοτιμές του πίνακα στο γράφημα Laplacian. Στη συνέχεια αντιπροσωπεύει σημεία σύμφωνα με τις σειρές του H . Λόγω των ιδιοτήτων του γραφήματος Laplacian, αυτή η αλλαγή αναπαράστασης είναι αρκετά χρήσιμη. Σε πολλές περιπτώσεις, αυτή η αλλαγή αναπαράστασης επιτρέπει στον απλό αλγόριθμο k-means να ανιχνεύει απρόσκοπτα τα συμπλέγματα. Διαισθητικά, εάν το H είναι όπως ορίζεται στο προαναφερθέν λήμμα, τότε κάθε σημείο στη νέα αναπαράσταση είναι ένα διάνυσμα δείκτη, του οποίου η τιμή είναι μη μηδενική, μόνο στο στοιχείο που αντιστοιχεί στο σύμπλεγμα στο οποίο ανήκει.

3.3. Αλγόριθμος DBSCAN

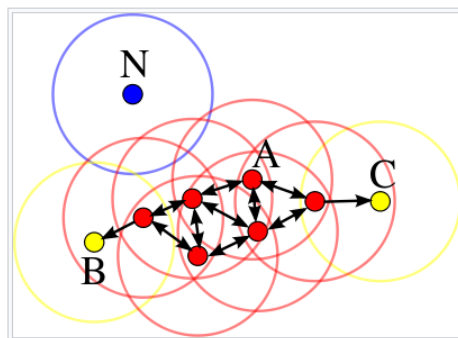
Η κύρια ιδέα πίσω από τον DBSCAN είναι ότι ένα σημείο ανήκει σε ένα σύμπλεγμα εάν είναι κοντά σε πολλά σημεία από αυτό το σύμπλεγμα. Υπάρχουν δύο βασικές παράμετροι του DBSCAN:

- ϵ : Η απόσταση που καθορίζει τις γειτονιές (neighborhoods). Δύο σημεία θεωρούνται γειτονικά εάν η απόσταση μεταξύ τους είναι μικρότερη ή ίση με ϵ .
- minPts : Ελάχιστος αριθμός σημείων δεδομένων για τον ορισμό ενός συμπλέγματος.

Με βάση αυτές τις δύο παραμέτρους, τα σημεία ταξινομούνται ως κεντρικό σημείο, συνοριακό σημείο ή εξωτερικό:

- Κεντρικό σημείο: Ένα σημείο είναι ένα κεντρικό σημείο εάν υπάρχουν τουλάχιστον minPts (συμπεριλαμβανομένου του ίδιου του σημείου) στη γύρω περιοχή με ακτίνα eps .
- Συνοριακό σημείο: Ένα σημείο είναι ένα συνοριακό σημείο εάν είναι προσβάσιμο από ένα κεντρικό σημείο και υπάρχουν λιγότεροι από minPts στη γύρω περιοχή.
- Εξωτερικό: Ένα σημείο είναι ένα εξωτερικό σημείο εάν δεν είναι κεντρικό και δεν είναι προσβάσιμο από κανένα κεντρικό σημείο.

Αυτά τα σημεία μπορεί να εξηγηθούν καλύτερα με οπτικοποιήσεις όπως φαίνεται και από το σχήμα 11 παρακάτω.



Σχήμα 11 - Παράδειγμα DBSCAN

Σε αυτή την περίπτωση, ο minPts είναι 4. Τα κόκκινα σημεία είναι κεντρικά σημεία επειδή υπάρχουν τουλάχιστον 4 σημεία εντός της γύρω περιοχής με ακτίνα eps . Αυτή η περιοχή εμφανίζεται με τους κύκλους στο σχήμα. Τα κίτρινα σημεία είναι συνοριακά επειδή είναι προσβάσιμα από ένα κεντρικό σημείο και έχουν λιγότερα από 4 σημεία στη γειτονιά τους. Προσβάσιμο σημαίνει να βρίσκεται στη γύρω περιοχή ενός κεντρικού σημείου. Τα σημεία B και Γ έχουν δύο σημεία (συμπεριλαμβανομένου του ίδιου του σημείου) μέσα στη γειτονιά τους (δηλαδή τη γύρω περιοχή με ακτίνα eps). Τέλος, το N είναι ένα εξωτερικό σημείο γιατί δεν είναι ένα κεντρικό σημείο και δεν μπορεί να επιτευχθεί από ένα κεντρικό σημείο. Στη συνέχεια, και έχοντας μάθει πλέον τους ορισμούς των παραμέτρων και τα διαφορετικά σημεία τύπου, γίνεται ανάλυση της λειτουργίας του αλγορίθμου με βάση τα παρακάτω:

- Τα minPts και eps είναι καθορισμένα.
- Ένα σημείο εκκίνησης επιλέγεται τυχαία στην περιοχή γειτονιάς του και προσδιορίζεται χρησιμοποιώντας την ακτίνα eps . Εάν υπάρχουν τουλάχιστον

minPts στη γειτονιά, το σημείο επισημαίνεται ως κεντρικό σημείο και ξεκινά ένας σχηματισμός συμπλέγματος. Εάν όχι, το σημείο επισημαίνεται ως θόρυβος. Μόλις ξεκινήσει ένας σχηματισμός συμπλέγματος (π.χ. σύμπλεγμα A), όλα τα σημεία μέσα στη γειτονιά του αρχικού σημείου γίνονται μέρος του συμπλέγματος A. Εάν αυτά τα νέα σημεία είναι επίσης κεντρικά σημεία, τα σημεία που βρίσκονται στη γειτονιά τους προστίθενται επίσης στο σύμπλεγμα A.

- Το επόμενο βήμα είναι να επιλεγεί τυχαία ένα άλλο σημείο μεταξύ των σημείων στα οποία δεν έχει γίνει επίσκεψη στα προηγούμενα βήματα. Τότε ισχύει η ίδια διαδικασία.
- Αυτή η διαδικασία ολοκληρώνεται όταν έχει γίνει επίσκεψη σε όλα τα σημεία.

Εφαρμόζοντας αυτά τα βήματα, ο αλγόριθμος DBSCAN είναι σε θέση να βρει περιοχές υψηλής πυκνότητας και να τις διαχωρίσει από περιοχές χαμηλής πυκνότητας. Ένα σύμπλεγμα περιλαμβάνει κεντρικά σημεία που είναι γειτονικά (δηλαδή προσβάσιμα το ένα από το άλλο) και όλα τα εξωτερικά σημεία αυτών των κεντρικών σημείων. Η απαραίτητη προϋπόθεση για να σχηματιστεί ένα σύμπλεγμα είναι να υπάρχει τουλάχιστον ένα κεντρικό σημείο. Αν και είναι πολύ απίθανο, υπάρχει η περίπτωση ένα σύμπλεγμα να αποτελείται από ένα μόνο κεντρικό σημείο και τα αντίστοιχα συνοριακά του σημεία.

3.4 Ιεραρχική συσταδοποίηση

Η Ιεραρχική Συσταδοποίηση παράγει ένα σύνολο από εμφωλευμένες συστάδες οργανωμένες σε ένα ιεραρχικό δέντρο. Μπορεί να παρασταθεί με ένα δενδρόγραμμα, το οποίο καταγράφει τις ακολουθίες από συγχωνεύσεις ή τους διαχωρισμούς των ομάδων, στις περιπτώσεις της συσσωρευτικής και διαιρετικής συσταδοποίησης αντίστοιχα. Η πρώτη αρχίζει με τα σημεία ως ξεχωριστές συστάδες ενώ η δεύτερη με μία συστάδα που περιέχει όλα τα σημεία. Η συσσωρευτική (Agglomerative) συγχωνεύει ομάδες με ένα κριτήριο απόστασης έως ότου φτάσει σε μια ομάδα ενώ η διαιρετική (Divisive) διαχωρίζει μία συστάδα με κάποιο κριτήριο διαχωρισμού, μέχρι να δημιουργηθούν τόσες συστάδες όσα και τα σημεία. Ο Agglomerative, ο οποίος και υλοποιείται στο πλαίσιο της παρούσας διπλωματικής, έχει τα εξής βήματα

1: Υπολογισμός του Πίνακα Γειτνίασης βάσει μιας δοσμένης μετρική απόστασης

2: Έστω κάθε σημείο αποτελεί και μια συστάδα

3: **Repeat**

4: Συγχώνευση των δύο κοντινότερων συστάδων βάσει ενός κριτηρίου

5: Ενημέρωση του Πίνακα Γειτνίασης

6: **Until** να μείνει μία μόνο συστάδα

Η απόσταση δύο σημείων υπολογίζεται συνήθως με χρήση της Ευκλείδειας απόστασης, ενώ το βήμα 4 γίνεται με διάφορα κριτήρια όπως

- **Μοναδικός σύνδεσμος:** Η ελάχιστη απόσταση ενός σημείου της συστάδας C_i από ένα σημείο της συστάδας C_j
- **Πλήρης σύνδεσμος:** Η μέγιστη απόσταση μεταξύ ενός σημείου της συστάδας C_i και ενός σημείου της συστάδας C_j
- **Μέσος όρος ομάδας :** Ο μέσος όρος της απόστασης ανά ζεύγη μεταξύ σημείων της συστάδας C_i και της συστάδας C_j

Κεφάλαιο 4. Βιβλιογραφική επισκόπηση μεθόδων συσταδοποίησης δεδομένων κίνησης

Μέρος Α – Παραδοσιακές μέθοδοι

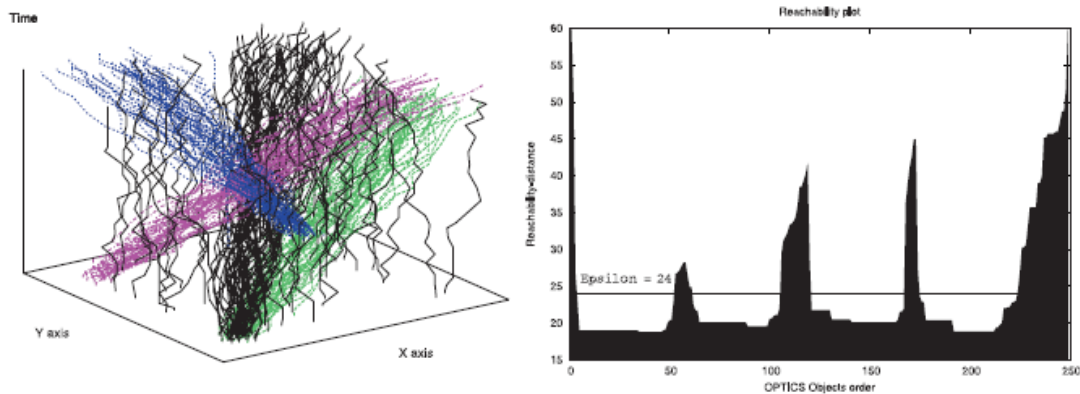
Σε αυτό το κεφάλαιο γίνεται βιβλιογραφική ανασκόπηση ερευνών που ασχολούνται με την χρήση τεχνικών ομαδοποίησης σε δεδομένα κίνησης. Η επισκόπηση έγινε με την εξέταση των ερευνών από την παλαιότερη χρονολογικά ώστε να γίνουν εμφανείς οι διαφορές και η εξέλιξη στις μεθόδους και στην αντίληψη του προβλήματος. Η παλαιότερη χρονολογικά έρευνα των Gaffney και Padhraic, (1999) εξέτασε αλγόριθμους συσταδοποίησης δεδομένων κίνησης θεωρώντας τα ως μια ομάδα μεταβλητών που εξαρτώνται από τον χρόνο. Οι μέθοδοι αυτοί βασίζονταν στο αλγόριθμο EM (Expectation – Maximization) που περιγράφει μια γενική διαδικασία την εύρεση παραμέτρων με την μέγιστη πιθανοφάνεια (likelihood) του υπό εξέταση μοντέλου. Πιο συγκεκριμένα, ο αλγόριθμος που εφάρμοσε βασιζόταν στον επανυπολογισμό των παραμέτρων θέσης σε ένα σταθμισμένο γραμμικό μοντέλο όπου τα βάρη αντιστοιχούσαν στην πιθανότητα θέσης. Η διαδικασία αυτή συνεχιζόταν μέχρι την εξαγωγή αποτελεσμάτων που συμφωνούσαν με τις οριακές τιμές (thresholds) που είχαν τεθεί. Οι εισαγόμενοι μέθοδοι επέτρεπαν τον ταυτόχρονο υπολογισμό περισσότερων της μιας τροχιάς και αναφερόταν σε τροχιές μίας ή δύο διαστάσεων. Επίσης, προέκτεινε την μελέτη του σε μη παραμετρικά δεδομένα, δηλαδή δεδομένα που δεν είχαν τις προϋποθέσεις για την εκτέλεση παλινδρόμησης, π.χ. κανονική κατανομή. Σύμφωνα με τους Gaffney και Padhraic (1999), η μέθοδος αυτή ήταν καλύτερη σε σύγκριση με τις μεθόδους SVM και είχε ιδιαίτερα καλά αποτελέσματα σε οπτικοακουστικές εφαρμογές και πιο συγκεκριμένα στην ροή εικόνας (video streaming).

Οι Nanni και Padreschi,(2006) εξέτασαν μεθόδους συσταδοποίησης δεδομένων κίνησης με δεδομένα που προέρχονται από την GSM κινητή τηλεφωνία. Βασική του αρχή ήταν η ποσοτικοποίηση της μέσης απόστασης μεταξύ δύο σημείων t_1 και t_2 . Αυτή έγινε με την βοήθεια της εξίσωσης

$$D(\tau_1, \tau_2)|_T = \frac{\int_T d(\tau_1(t), \tau_2(t))dt}{|T|} \quad (E3.1)$$

Όπου $\tau_i(t)$, $i=1,2$ είναι η θέση του σημείου τ_i στον χρόνο t , T το διάστημα μεταξύ των σημείων και d η Ευκλείδεια απόσταση μεταξύ των σημείων. Ο αλγόριθμος που χρησιμοποιήσαν ονομάζεται OPTICS (Ordering points to identify the clustering structure) των (Ankerst et al., 1999), ο οποίος χρησιμοποιείται ευρέως για την συσταδοποίηση δεδομένων και βασίζεται στην πυκνότητα των παρατηρήσεων (density based clustering).

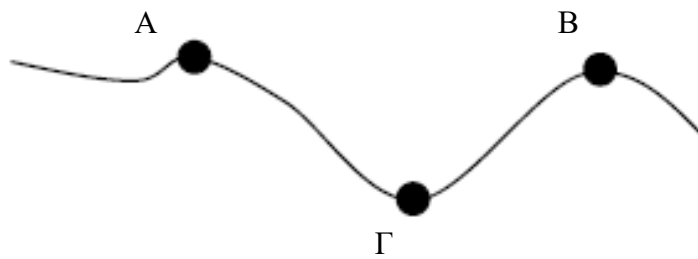
Τα δεδομένα που χρησιμοποιήθηκαν για την εξέταση της απόδοσης των μεθόδων ήταν προσομοιωμένα (simulated) με την βοήθεια του αλγόριθμου C4C, μιας νεότερης έκδοσης του αλγορίθμου CENTRE (CELLular Network Trajectories Reconstruction Environment) των Giannotti et al., 2005. Τα αποτελέσματα της εφαρμογής της μεθόδου έδειξαν πολύ καλή συμπεριφορά και ακριβή αποτελέσματα όπως φαίνεται και στην εικόνα 12.



Σχήμα 12 - Αρχικά δεδομένα (αριστερά) και το εξαγόμενο διάγραμμα εγγύτητας (reachability plot με την βοήθεια του αλγόριθμου OPTICS (δεξιά),

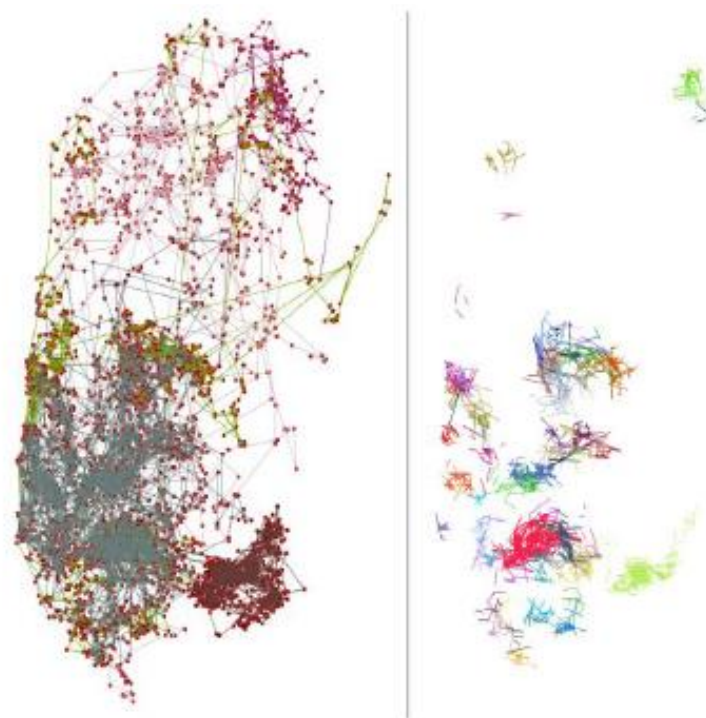
Πηγή: Nanni, & Pedreschi, 2006

Οι Chang & Chom, (2009), προτείνουν μια προσέγγιση για την οπτικοποίηση της συσταδοποίησης της τροχιάς των δεδομένων κίνησης με βάση τις συστάδες υπό-τροχιάς από τα δεδομένα αρχικής κίνησης μεγάλης κλίμακας. Η μέθοδος αυτή βασίζεται στην υποδιαίρεση των αρχικών δεδομένων σε ένα μικρότερο σύνολο δεδομένων κίνησης με την εύρεση των σημείων μεταβολής της κίνησης (εικόνα 13).



Σχήμα 13 - Παράδειγμα τροχιάς με 3 σημεία μεταβολής της κίνησης A,B και Γ.

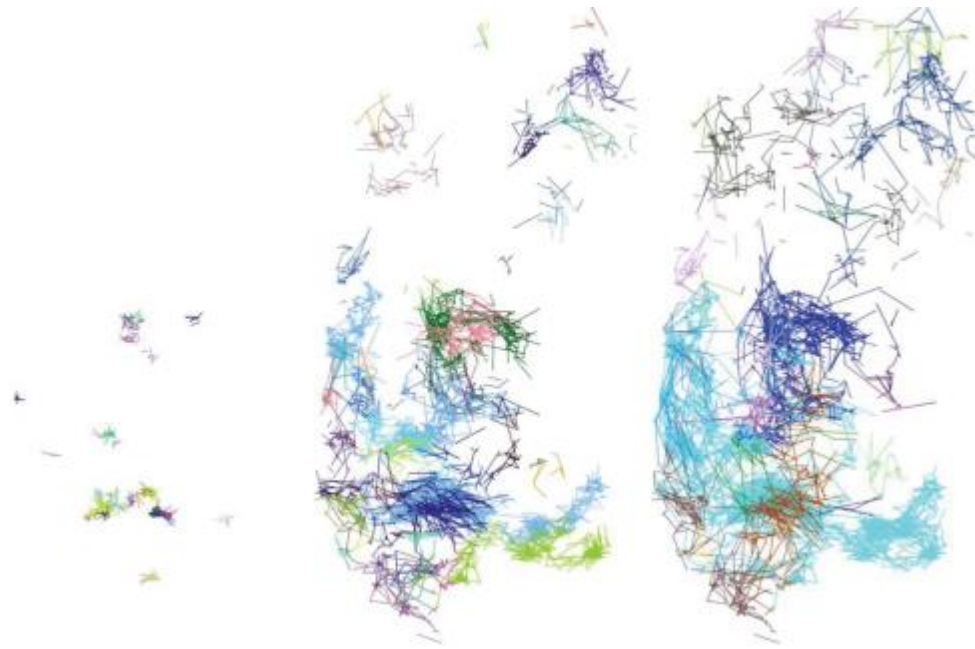
Στην συνέχεια έγιναν υπολογισμοί της απόστασης Fréchet³ για τον υπολογισμό της ομοιότητας μεταξύ των υποσυνόλων και την χρήση μεθόδων συσταδοποίησης με βάση την πυκνότητα τους. Τα αποτελέσματα έδειξαν μεγάλη ακρίβεια στην συσταδοποίηση των υπό-ομάδων (εικόνα 14). Επιπλέον η συσταδοποίηση αυτή έχει το πλεονέκτημα της παραμετροποίησης, κάτι που επιτρέπει την επιλογή της τάξης μεγέθους της απόστασης μεταξύ των κέντρων (συντελεστής πυκνότητας).



Σχήμα 14 - Αρχικά δεδομένα (αριστερά) και συσταδοποίηση υπό-ομάδων (δεξιά).

Πηγή: Chang & Chom, 2009 και ίδια επεξεργασία.

³ Το μέγιστο κάτω άκρο μιας ομάδας καμπυλών κίνησης σε χρόνο t.



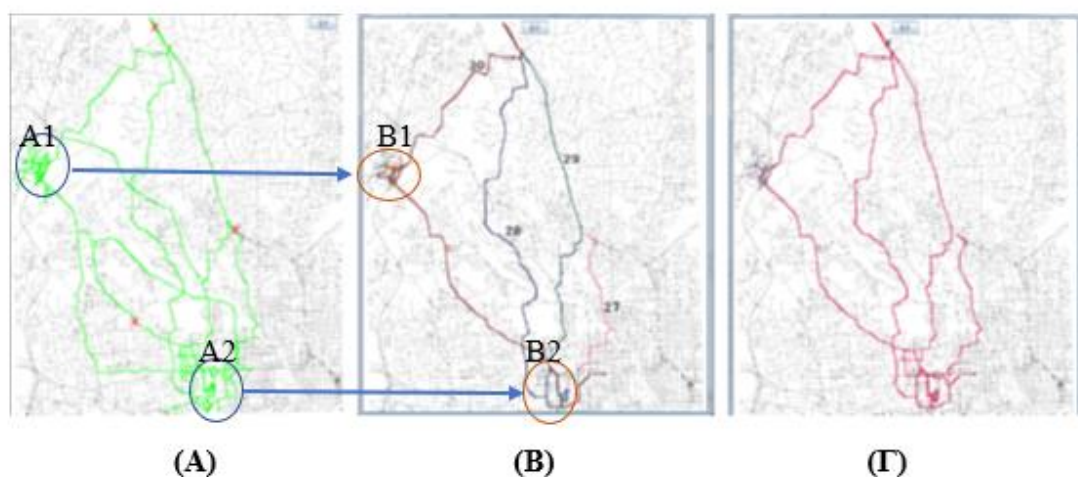
Σχήμα 15 - Εφαρμογή παραμέτρων συσταδοποίησης με βάση την πυκνότητα (Α. Συντελεστής πυκνότητας =10, Β. Συντελεστής πυκνότητας =31 και Γ. Συντελεστής πυκνότητας = 47).

Πηγή: Chang & Chom, (2009)

Οι Hung, Liu και Omnicinski, (2012), υποστηρίζουν στην έρευνα τους ότι όταν η χρησιμότητα της χωρικής ομαδοποίησης δεδομένων κίνησης στοχεύει σε οδικές εφαρμογές, τότε ο υπολογισμός της Ευκλείδειας απόστασης δεν αποτελεί πλέον ένα επαρκές μέτρο. Αυτό συμβαίνει διότι σε αυτές τις περιπτώσεις υπεισέρχονται οι παράμετροι του ρυθμού και του όγκου της κυκλοφορίας του οδικού δικτύου. Η προτεινόμενη μέθοδος συσταδοποίησης των Hung, Liu και Omnicinski, (2012), ονομάζεται NEAT (Network Aware Trajectory).

Η μέθοδος αυτή υπολογίζει και άλλες παραμέτρους κατά την συσταδοποίηση των δεδομένων κίνησης. Οι παράμετροι αυτοί είναι οι φυσικοί περιορισμοί του οδικού δικτύου, η εγγύτητα του δικτύου και η ροή της κυκλοφορίας μεταξύ διαδοχικών οδικών τμημάτων. Οι συστάδες που ανακαλύφθηκαν από την μέθοδο NEAT είναι υποομάδες δεδομένων κίνησης που περιγράφουν τόσο την πυκνή όσο και την έντονη ροή της κυκλοφορίας κινούμενων αντικειμένων. Τα πειραματικά μας αποτελέσματα έδειξαν ότι η προσέγγιση NEAT είναι πολύ ακριβής και ταχύτερη, συγκρινόμενη με παλαιότερες μεθόδους.

Τα αποτελέσματα της εφαρμογής της μεθόδου NEAT παρουσιάζονται στην εικόνα 16. Σύμφωνα με αυτά είναι φανερή η αφαίρεση του θορύβου (σχήμα Β) αλλά και η βελτιστοποίηση (εδώ μείωση) του αριθμού των συστάδων (σχήμα Γ).

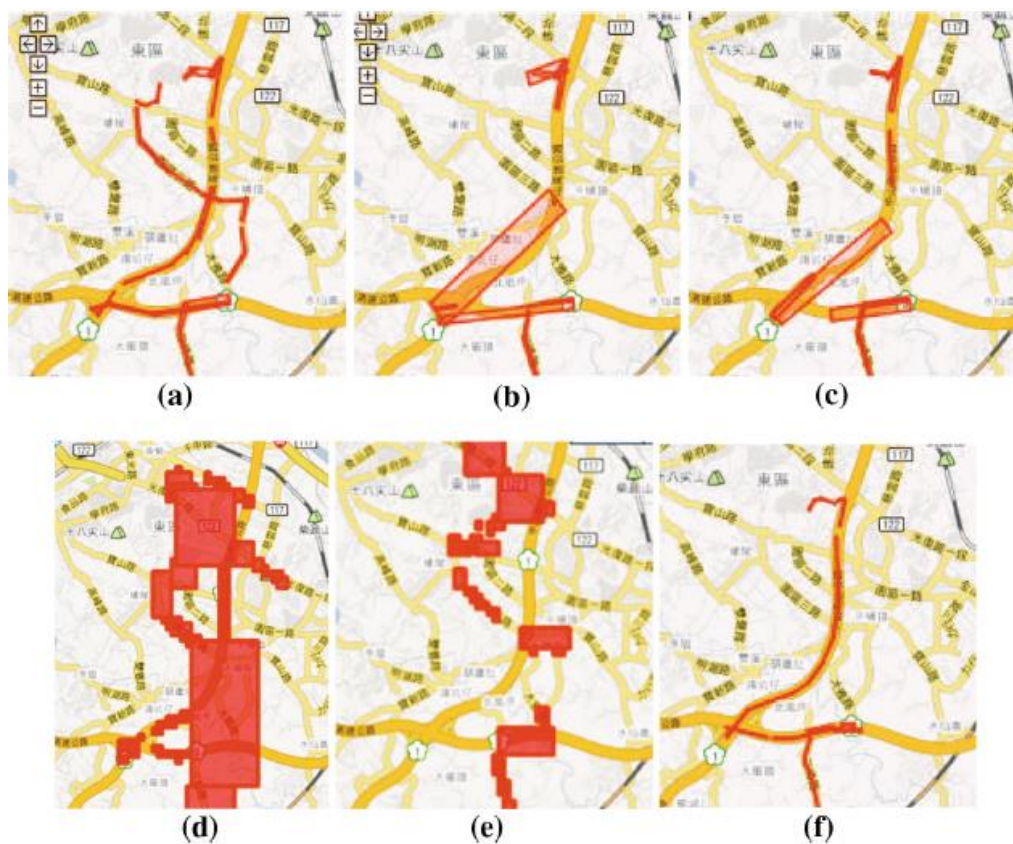


Σχήμα 16 - Αρχικά δεδομένα (Α), Αρχική συσταδοποίηση (Β) και Βέλτιστη συσταδοποίηση (Γ) με την μέθοδο NEAT
Πηγή: Hung, Liu & Omnicinski, 2012 & Ιδία επεξεργασία

Οι Hung, Peng και Lee (2015), πρότειναν ένα νέο πλαίσιο εξόρυξης μοτίβων της τροχιάς των δεδομένων κίνησης, το Clustering and Aggregating Clues of Trajectories (CACT), για την ανακάλυψη σχεδίων τροχιάς που αντιπροσωπεύουν τις συχνότερες συμπεριφορές ενός χρήστη. Εκτός από τις χωρικές και χρονικές μεροληψίες π.χ. σφάλματα μέτρησης, παρατήρησαν ότι οι τροχιές περιέχουν περιόδους αδράνειας δηλ. χρονικές περιόδους όπου δεν υπάρχουν δεδομένα για να περιγράψουν τις κινήσεις των χρηστών, κάτι που περιεγράφηκε στις δυσκολίες συλλογής δεδομένων στο πρώτο κεφάλαιο, οι οποίες αποτελούν πρόκληση στην εξόρυξη δεδομένων κίνησης.

Σύμφωνα με τους Hung, Peng και Lee, (2015), μια συμπεριφορά κίνησης θα αφήσει κάποιες ενδείξεις στις διάφορες πορείες του δείγματος-παρατήρησης. Αυτές οι ενδείξεις μπορούν να εξαχθούν από χωρικά και χρονικά συντεταγμένα σημεία δεδομένων από τις παρατηρούμενες τροχιές. Με βάση αυτήν την παρατήρηση, πρότειναν την εξέταση των ομοιοτήτων που παρουσιάζουν οι τροχιές της κίνησης με τη μέτρηση των ενδείξεων μεταξύ δύο τροχιών και την ανάλογη συσταδοποίησή τους.

Ο αλγόριθμος συσταδοποίησης που κατασκεύασαν με βάση αυτό το σκεπτικό ονομάστηκε CACT (Clue-Aware Trajectory Aggregation). Ο αλγόριθμος αυτός εφαρμόστηκε σε πειράματα χρησιμοποιώντας συνθετικά και πραγματικά σύνολα δεδομένων. Παρουσιάζει ομοιότητες με το σκεπτικό των Chang και Chom, (2009) μόνο που το κριτήριο συσταδοποίησης τώρα είναι η ομοιότητα μεταξύ των τροχιών κίνησης και όχι η μεταβολή της κίνησης. Τα αποτελέσματα εφαρμογής του αλγορίθμου CACT παρουσιάζεται στην εικόνα 17 και έδειξαν ότι ο CACT είναι πιο αποτελεσματικός στην ανακάλυψη μοτίβων τροχιάς από πιο προηγμένες τεχνικές εξόρυξης μοτίβων τροχιάς όπως ο TPM (Cao et al, 2019) και SFP (Cao, Mamoulis και Cheung, 2005).



Σχήμα 17 - Αποτελέσματα εφαρμογής αλγορίθμων (a) Πραγματική τροχιά, (b) SFP-L, (c) SFP-C, (d) TPM-L, (e) TPM-C, (f) CACT

Πηγή: Hung, Peng & Lee, 2015.

Οι Xu et al., (2015), πρότειναν ένα πλαίσιο που βασίζεται στην παροχή συγκεκριμένης πληροφορίας (συλλογής δεδομένων) για την συσταδοποίηση δεδομένων κίνησης. Αντιμετωπίζοντας τις προκλήσεις της ομαδοποίησης τροχιάς, π.χ., μεγάλη μεταβλητότητα της συστάδας και αμφισημίες μεταξύ των συστάδων, εισήγαγαν μια προσαρμοσμένη διαδικασία εκτίμησης των συστάδων με κεντρικά σημεία ή πυρήνες

(kernels) σε αντίθεση με τις προηγούμενες μεθόδους που εξέταζαν την πυκνότητα των συστάδων. Αυτού του είδους η εκτίμηση αξιοποιεί αποτελεσματικά και τις πολλαπλές δομικές πληροφορίες σε μια τροχιά και τα τοπικά μοτίβα κίνησης σε πολλές τροχιές, έτσι ώστε η διάκριση του κεντρικού σημείου συσταδοποίησης να αποτελεί ακριβές μέτρο συσταδοποίησης. Η διαδικασία αυτή βελτιστοποιείται με την εξέταση επιπλέον παραμέτρων κίνησης όπως η μεταβλητή ταχύτητα, η οποία χρησιμοποιεί τις εκτιμώμενες ταχύτητες για τη ρύθμιση των βέλτιστων κεντρικών σημείων, έτσι ώστε να διασφαλίζεται η ομαλότητα στην μετάβαση του τελικού σημείου.

Χρησιμοποιώντας αυτή την προσέγγιση, μειώνεται η μεταβλητότητα των συστάδων και ορίζονται πιο ευδιάκριτα τα όρια μεταξύ τους. Τα πειραματικά αποτελέσματα έδειξαν ότι η προσέγγισή μας είναι ανώτερη από της τελευταίας τεχνολογίας προσεγγίσεις τόσο σε ακρίβεια αποτελεσμάτων όσο και ορθότητα εφαρμογής της (robustness). Επιπλέον η μέθοδος επιτρέπει και την εξαγωγή αποτελεσμάτων ανάλυσης της διαδρομής (Route analysis: Εικόνα 18)



Σχήμα 18 - Εφαρμογή μεθόδου των Xu et al., (2015), σε δεδομένα ανθρώπινης κίνησης (A: Αρχικά δεδομένα, B: Συστάδες, Γ: Route Analysis)

Πρόσφατες έρευνες έχουν επικεντρωθεί στην ομαδοποίηση σε υποτροχιές (sub trajectory) όπου μια υποτροχιά $r_{i,j}$ είναι το υποσύνολο $\{r_1, r_2, \dots, r_j\}$ της τροχιάς r που συμβολίζει την κίνηση ενός αντικειμένου μεταξύ των χρονικών στιγμών t_i και t_j ($i < j$) όπου $i, j \in 1, \dots, N$. Ενδεικτική είναι η εργασία από τους Tampakis et al., (2019) με σκοπό την αναγνώριση συστάδων ολοκληρωμένων τροχιών καθώς, όπως αναφέρουν, υπάρχουν ατέλειες στην αναγνώριση ολοκληρωμένων τροχιών μικρής διάρκειας οι οποίες συχνά παραλείπονται από τους αλγόριθμους συσταδοποίησης. Η έρευνα τους αφορούσε κατανεμημένη συσταδοποίηση υπό-τροχιών και η μέθοδος που πρότειναν ονομάστηκε Scalable Distributed Subtrajectory Clustering. Βασικό πρόβλημα κατά την αναγνώριση των υποτροχιών είναι ότι οι υποτροχιές που θα συσταδοποιηθούν δεν είναι γνωστές εξαρχής αλλά θα πρέπει να αναγνωριστούν δυναμικά με βάση γειτονικές

υποτροχιές. Η επίλυση του προβλήματος έγινε με τον διαχωρισμό στα εξής υπό προβλήματα (1) ένωση των υποτροχιών (Sub trajectory Join), (2) διαχωρισμός των τροχιών (Trajectory Segmentation), (3) συσταδοποίηση των υποτροχιών και ανίχνευση ακραίων τιμών (Sub trajectory Clustering and Outlier Detection) και (4) συσταδοποίηση κατανεμημένων υποτροχιών.

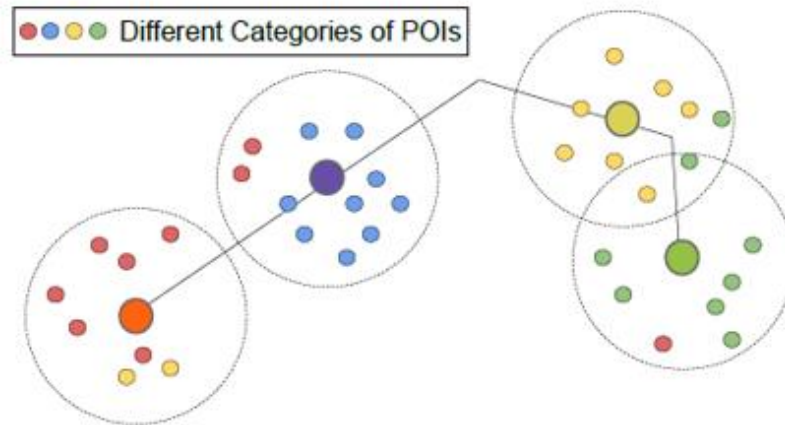
Η επισκόπηση των προηγούμενων ερευνών έδειξε ότι οι εφαρμοζόμενες μέθοδοι ποικίλουν ανάλογα με το πρόβλημα που εξετάζουν, το πεδίο εφαρμογής τους αλλά και τις μεθόδους που προσπάθησαν να βελτιώσουν. Σχεδόν όλες οι μέθοδοι αυτοί βασιζόταν στην συσταδοποίηση με βάση την πυκνότητα των δεδομένων (σημείων) και όλες ανέφεραν βελτίωση της αποτελεσματικότητας σε σύγκριση με προηγούμενες μεθόδους. Το γεγονός αυτό δείχνει ότι οι μέθοδοι συσταδοποίησης δεδομένων κίνησης αυξάνονται τα τελευταία χρόνια (Pelekis et al., 2011; Bian et al., 2018; Feng & Zhu, 2016; Tampakis et al., 2019) και ότι αποτελεί πλέον ένα κλάδο της επιστήμης που αν και αναφέρεται κυρίως σε εμπορικές εφαρμογές αποτελεί βασικό αντικείμενο ενδιαφέροντος της ακαδημαϊκής κοινότητας.

Μέρος Β – Νεότερες μέθοδοι βασισμένες σε βαθιά μάθηση (DNN)

Στο δεύτερο μέρος αυτού του κεφαλαίου γίνεται η παρουσίαση πιο πρόσφατων μεθόδων τεχνικών ομαδοποίησης σε δεδομένα κίνησης, οι οποίες ενσωματώνουν τη γνώση από τεχνικές βαθιάς μάθησης. Η χρησιμοποίηση της βαθιάς μάθησης συνήθως αφορά την εξαγωγή χαρακτηριστικών από τα δεδομένα ή κάποιου είδους προ επεξεργασίας του αρχικού σετ δεδομένων ή κάποιου είδους μετασχηματισμό στο σετ δεδομένων (όπως πχ. το embedding).

Η έρευνα των Yue et al., (2019) αφορούσε την συσταδοποίηση τροχιών βελτιώνοντας τα συχνά προβλήματα παρόμοιων αλγορίθμων συσταδοποίησης, που παρουσιάζονται στην περίπτωση τροχιών που μοιάζουν μεταξύ τους, δηλαδή παρουσιάζουν παρόμοιες συμπεριφορές κίνησης, όπως και η υψηλή χρήση πόρων που απαιτούν για την εφαρμογή τους. Για την αντιμετώπιση αυτών των προκλήσεων, πρότειναν μια μη-επιβλεπόμενης μάθησης μεθοδολογία βασισμένη στα νευρωνικά δίκτυα για την ομαδοποίηση συμπεριφοράς της κίνησης. Ο αλγόριθμος που ονομάζεται Deep Embedded Trajectory Clustering network (DETECT) περιέχει τρία στάδια: πρώτα μετασχηματίζει τις αρχικές τροχιές συνοψίζοντας τα κρίσιμα σημεία τους και τα

ενδυναμώνουν μέσω της γεωγραφικής τους θέσης π.χ., χρησιμοποιώντας σημεία ενδιαφέροντος (Points Of Interest ή POI) (Σχήμα 19).



(α)



(β)

Σχήμα 19 - Σημεία ενδιαφέροντος (α) και εφαρμογή τους σε πραγματικά δεδομένα (β) (Πηγή: Yue et al., 2019).

Στο δεύτερο στάδιο, γίνεται η εκμάθηση των συνήθων αναπαραστάσεων η οποία έχει ήδη διαγραφεί, επιτρέποντας έτσι την εφαρμογή μεθόδων όπως ο k-means. Το τελευταίο στάδιο αφορά την άμεση εφαρμογή του αλγορίθμου για την βελτίωση της

συσταδοποίησης. Πιο συγκεκριμένα, η απόσταση μεταξύ δύο σημείων ενδιαφέροντος P και Q υπολογίζεται από τον τύπο

$$\ell_c = KL(P \parallel Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

όπου

$$p_{ij} = \frac{q_{ij}^2 / \sum_{i'} q_{i'j}}{\sum_{j'} (q_{ij'}^2 / \sum_{i'} q_{i'j'})}$$

και

$$q_{ij} = \frac{(1 + \|z_i - \mu_j\|^2)^{-1}}{\sum_{j'} (1 + \|z_i - \mu_{j'}\|^2)^{-1}}$$

που εκφράζει την πιθανότητα ανάθεσης της τροχιάς z_i στην τροχιά μ_j . Η εφαρμογή του αλγορίθμου DETECT με την βοήθεια των POI και των βελτιώσεων των Yue et al., (2019) έδειξαν βελτίωση κατά 40%.

Η εργασία των Harmening et al. (2020) προτείνουν ένα μοντέλο βασισμένο σε Τεχνητά Νευρωνικά δίκτυα Συνελίξεων (CNN) και RNN σε δεδομένα πραγματικού κόσμου τα οποία αφορούν την κίνηση σε έναν αυτοκινητόδρομο. Δημιουργούν έναν λανθάνων χώρο (latent space) ο οποίος χρησιμοποιείται για να ανακτήσει παρόμοιες σκηνές από το αρχικό σύνολο δεδομένων έτσι ώστε να αυξηθεί το dataset για να ενισχύσει το βήμα της εκπαίδευσης του ταξινομητή. Στη συνέχεια ομαδοποιούν τις διάφορες τροχιές κίνησης χρησιμοποιώντας ιεραρχικό αλγόριθμο ομαδοποίησης, εντοπίζοντας τέσσερα cluster τα οποία αφορούν περιπτώσεις τις κίνησης στο δρόμο όπως «πυκνή κίνηση» και «σχεδόν άδειος δρόμος». Η αποτελεσματικότητα της μεθοδολογίας αυτής έγκειται κυρίως στην χρήση των CNN, που αποτελούν την εξέλιξη των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) καθώς ενσωματώνουν στη δομή τους όχι μόνο τους τυπικούς νευρώνες των ΤΝΔ αλλά τελεστές περισσότερο σύνθετους (με βασικότερο τελεστή την πράξη της συνελίξης), που μπορούν να επιλύουν συνθετότερα προβλήματα. Τα CNN οργανώνονται σε πολλαπλά επίπεδα στρωμάτων (layers) συνδυάζοντας του τυπικούς νευρώνες των ΤΝΔ με τα μητρώα χαρακτηριστικών, την νέα δομή που ενσωματώνουν τα CNN. Δημιουργούν και απομνημονεύουν εμπειρική γνώση την οποία θα μπορούν να την χρησιμοποιήσουν στο μέλλον μέσω των ομάδων βαρών. Ο τρόπος απόκτησης

της εμπειρικής γνώσης προσομοιώνεται με τον τρόπο απόκτησης της γνώσης από τον ίδιο τον άνθρωπο μέσω ενός μηχανισμού μάθησης, που στην Υπολογιστική Νοημοσύνη είναι ο αλγόριθμος εκπαίδευσης.

Πιο συγκεκριμένα, έχοντας ως βάση το τρισδιάστατο πλέγμα στοιβαγμένης ταχύτητας (3D Stacked Velocity Grid) που παρουσιάζεται στη μελέτη των συγγραφέων (Gruner et al., 2017), ανέπτυξαν μια εκτεταμένη αναπαράσταση πλέγματος για την καλύτερη καταγραφή των χρονικών χαρακτηριστικών. Για να αποφύγουν τον περιορισμό σε δύο χωρικές διαστάσεις, αλλά και τη διάσταση του καναλιού, εισήγαγαν τον άξονα χρόνου, κατά μήκος του οποίου ευθυγραμμίζονται τα πλαίσια. Κάθε πλαίσιο σε ένα σημείο, αποτελείται από ένα δυαδικό κανάλι κατάληψης, το οποίο κωδικοποιεί τη θέση κίνησης των συμμετεχόντων σε ένα προκαθορισμένο σενάριο. Το αυτοματοποιημένο όχημα είναι στερεωμένο σε ένα κεντρικό κελί πλέγματος (σταθερή πλευρική και διαμήκης μετατόπιση) σε όλα τα σενάρια, επομένως οι κινήσεις των συμμετεχόντων στην κυκλοφορία είναι πάντα σχετικές με το αυτοματοποιημένο όχημα. Μπορούν επίσης να προστεθούν επιπλέον κανάλια που κωδικοποιούν περαιτέρω λειτουργίες στην αντίστοιχη τοποθεσία.

Στη συνέχεια και για την κωδικοποίηση του τετραδιάστατου πλέγματος εισόδου X , χρησιμοποιούνται τρισδιάστατες συνελίξεις που συνελίσσονται σε δύο χωρικές διαστάσεις μαζί με τη χρονική διάσταση. Οι συγγραφείς (Karpathy et al., 2014) και (Lea et al., 2016) έχουν ήδη χρησιμοποιήσει σε μελέτες τους τα τρισδιάστατα συνελκτικά νευρωνικά δίκτυα (convolutional neural networks - CNNs). Οι πρώτοι, για ταξινόμηση βίντεο μεγάλης κλίμακας και οι δεύτεροι, κατά την τμηματοποίηση βίντεο. Μετά από κάθε συνέλιξη στον κωδικοποιητή, εφαρμόζονται τα στρώματα της ομαλοποίησης παρτίδας και της μέγιστης συγκέντρωσης. Έπειτα, ο αποκωδικοποιητής κάνει χρήση των στρωμάτων άνω συνελίξεων και αποκατάστασης συγκέντρωσης, τα οποία εισήχθησαν στο δίκτυο αποσυνέλιξης (deconvolution network – DeconvNet) με σκοπό να αποκωδικοποιήσουν τη λανθάνουσα αναπαράσταση πίσω στη διάσταση εισόδου. Τέλος, το τελικό διάνυσμα συμφόρησης Z , κωδικοποιεί το σενάριο στη λανθάνουσα αναπαράσταση και χρησιμοποιείται σε επόμενο στάδιο για την εργασία ομαδοποίησης. Δεδομένου ότι η ταχύτητα και η επιτάχυνση είναι συνεχείς, οι συγγραφείς (Harmening et al., 2020) χρησιμοποιούν το μέσο τετράγωνο σφάλμα (Mean squared error) μεταξύ του πλέγματος στόχου X και του ανακατασκευασμένου πλέγματος \hat{X} ως απώλεια ανακατασκευής:

$$\mathcal{L}(\mathbf{X}, \hat{\mathbf{X}}) = \sum_{a=d_t} \sum_{b=d_x} \sum_{c=d} \sum_{d=n_e} (x_{a,b,c,d} - \hat{x}_{a,b,c,d})^2$$

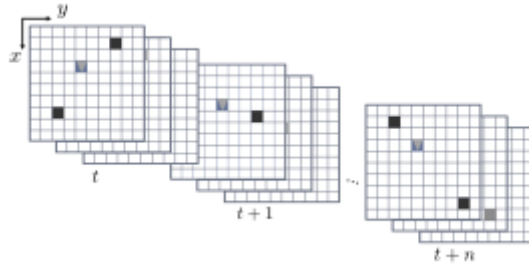
Το τελευταίο βήμα είναι η ανακατασκευή των συνόλων με βάση τα ανακατασκευασμένα διανύσματα ενσωμάτωσης $\hat{\mathbf{e}}_t$. Για το σκοπό αυτό, οι συγγραφείς (Harmening et al., 2020) χρησιμοποιούν το δίκτυο πρόβλεψης βαθιών συνόλων (Deep Set Prediction Network – DSPN), το οποίο μπορεί να θεωρηθεί ως προγνωστικός δείκτης συνόλων, με βάση ένα διάνυσμα ενσωμάτωσης. Όσον αφορά την απώλεια, οι συγγραφείς (Zhang et al., 2019) προτείνουν τη χρήση της απόστασης Chamfer $O(n^2)$ ή του ουγγρικού αλγόριθμου $O(n^3)$ για τον προσδιορισμό της απόστασης \mathcal{L}_{set} μεταξύ της ακολουθίας εισόδου των συνόλων \mathbf{X} και των ανακατασκευασμένων $\hat{\mathbf{X}}$, η οποία πολλαπλασιάζεται με την προβλεπόμενη επικάλυψη. Για παράδειγμα, η απόσταση Chamfer ταιριάζει τις πλησιέστερες περιπτώσεις από το \mathbf{X} στο $\hat{\mathbf{X}}$ και αντίστροφα:

$$\begin{aligned} \mathcal{L}_{\text{set}}(\hat{\mathbf{X}}, \mathbf{X}) &= \sum_i \min_j \|\hat{\mathbf{x}}_i - \mathbf{x}_j\|^2 \\ &+ \sum_j \min_i \|\hat{\mathbf{x}}_i - \mathbf{x}_j\|^2 \end{aligned}$$

Επιπλέον, για να σταθεροποιήσουν περαιτέρω τη διαδικασία εκπαίδευσης του επαναλαμβανόμενου νευρωνικού δικτύου (Recurrent neural network – RNN), οι συγγραφείς (Harmening et al., 2020) προσθέτουν το μέσο τετράγωνο σφάλμα μεταξύ των ενσωματώσεων και της ανακατασκευασμένης ενσωμάτωσης, το οποίο αποφέρει την τελική απώλεια:

$$\mathcal{L}_{\text{dspm}} = \mathcal{L}_{\text{set}} + \lambda \sum_{t=1}^T \mathcal{L}_{\text{mse}}(\mathbf{e}_t, \hat{\mathbf{e}}_t)$$

όπου το λ είναι μια παράμετρος στάθμισης. Πρόσθετες υπερπαραμέτροι στο δίκτυο πρόβλεψης διαδοχικών βαθιών συνόλων (Sequential Deep Set Prediction Network – SeqDSPN), είναι ο εσωτερικός ρυθμός εκμάθησης και ο αριθμός των εσωτερικών βημάτων βελτιστοποίησης του DSPN.



Σχήμα 20 Χωροχρονικό πλέγμα (Πηγή: Harmening et al., 2020)

Η εργασία των Boyle et al. (2017) παρουσιάζει μια βαθιά μάθησης προσέγγιση αναπαράστασης χαρακτηριστικών τροχιάς για να ενισχύσει την διαδικασία της ομαδοποίησης της τροχιάς και να αναγνωρίσει διάφορα πρότυπα κίνησης από βίντεο. Χρησιμοποιούν έναν autoencoder στον οποίο απομονώνουν την έξοδο του μικρότερου κρυφού νευρώνα του autoencoder για να εξάγουν χαρακτηριστικά από το σετ δεδομένων. Στην ουσία κάνουν μείωση διαστατικότητας κρατώντας τα χαρακτηριστικά αυτά τα οποία βρίσκονται στο κέντρο του autoencoder και μας μειώνουν αρκετά την υψηλή διαστατικότητα του σετ δεδομένων μας. Στο επόμενο βήμα, ο αλγόριθμος ομαδοποίησης τροχιάς εφαρμόζεται στο σετ δεδομένων με αυτά τα χαρακτηριστικά. Τα πρότυπα κίνησης εξάγονται με βάση την ελαχιστοποίηση της απόστασης από τα κέντρα των συστάδων. Η μεθοδολογία αυτή είναι αρκετά υποσχόμενα αποτελέσματα και δείχνει την δυναμική της μείωσης της διαστατικότητας και πόσο αυτή μπορεί να ενισχύσει την ομαδοποίηση του σετ δεδομένων.

Πιο συγκεκριμένα, προκειμένου να δημιουργηθεί το διάνυσμα χαρακτηριστικών \mathbf{f}_j για μια τροχιά \mathcal{X}_j , εκπαιδεύεται σε πρώτη φάση ένας αυτόματος κωδικοποιητής (επίσης γνωστός δίκτυο Diabolo) με σκοπό να αναπαράγει το σύνολο των τροχιών \mathcal{X} . Μόλις εκπαιδευτεί ένα δίκτυο, η έξοδος του μικρότερου επιπέδου του χρησιμοποιείται ως το διάνυσμα χαρακτηριστικών \mathbf{f}_j . Επομένως, ένα ξεχωριστό δίκτυο πρέπει να εκπαιδεύεται για κάθε σύνολο δεδομένων, λόγω του ποικίλου μήκους των διανυσμάτων εισόδου. Ένα δίκτυο που αποτελείται μόνο από πλήρως συνδεδεμένα επίπεδα, δεν μπορεί να χειριστεί διανύσματα εισόδου διαφορετικών μεγεθών, χωρίς την ανάγκη να εισάγει περαιτέρω μεθόδους κανονικοποίησης. Έτσι, λόγω αυτού και της διαφοράς στα σημεία διαφορετικών συνόλων δεδομένων, επιλέχθηκε να χρησιμοποιηθούν ξεχωριστά δίκτυα. Οι έξοδοι οποιουδήποτε από τα επίπεδα σε ένα εκπαιδευμένο δίκτυο, μπορούν να χρησιμοποιηθούν ως αναπαράσταση της εισόδου

λόγω της ικανότητας του υπόλοιπου δικτύου να αναπαράγει το διάνυσμα εισόδου. Σε αυτήν την περίπτωση, η είσοδος είναι η διανυσματικοποίηση των δεδομένων του \mathbb{X}_j και συμβολίζεται ως $\mathbf{V}_{\mathbb{X}_j}$:

$$\mathbf{V}_{\mathbb{X}_j} = [x_{k,j}, y_{k,j}]_{k=k_{\text{start}}^j}^{k_{\text{end}}^j}$$

Ένα νευρωνικό δίκτυο είναι ένας συνδυασμός μικρών μονάδων που ονομάζονται νευρώνες, οι οποίοι είναι δομημένοι σε πολλαπλά στρώματα. Ο τύπος του νευρώνα που χρησιμοποιείται σε αυτή τη μελέτη των συγγραφέων (Boyle et al., 2017), βασίζεται στον νευρώνα των McCullough-Pitts. Αυτός, πολλαπλασιάζει ένα μονοδιάστατο διάνυσμα εισόδου με ένα διάνυσμα βάρους που στη συνέχεια αθροίζεται με έναν κόμβο πόλωσης και τελικώς εξάγει μια μοναδική τιμή:

$$y_{l,n} = \sum_i^{I_l} (w_{i,n} \mathbf{x}_{l,i}) + b_n$$

όπου το $y_{l,n}$ προσφέρει την έξοδο προενεργοποίησης-συνάρτησης y του νευρώνα n , στο στρώμα l . Το I_l είναι το μήκος του διανύσματος εισόδου X_l για ένα συγκεκριμένο στρώμα, το w_i και το $\mathbf{x}_{l,i}$ είναι μια τιμή στο διάνυσμα βάρους και στο διάνυσμα εισόδου αντίστοιχα, και το b είναι η τιμή πόλωσης. Έπειτα, το διάνυσμα βάρους αρχικοποιείται σε τυχαίες τιμές και χρησιμοποιείται μια σιγμοειδής συνάρτηση για τη συνάρτηση ενεργοποίησης:

$$Y_{l,n} = \frac{1}{1 + e^{-y_{l,n}}}$$

όπου Y είναι η έξοδος του νευρώνα. Στη συνέχεια, οι νευρώνες τοποθετούνται ο ένας δίπλα στον άλλο ως στρώμα. Η έξοδος ενός στρώματος l μπορεί να περιγραφεί από την ακόλουθη εξίσωση:

$$L_l = [Y_{l,1}, \dots, Y_{l,L_l}]$$

όπου το στρώμα L_l είναι ένα διάνυσμα που περιέχει όλες τις εξόδους των νευρώνων στο στρώμα. Μετά, παρέχεται το επόμενο επίπεδο, το οποίο έχει ως διάνυσμα εισόδου του (γνωστό ως πλήρως συνδεδεμένο στρώμα), την έξοδο του προηγούμενου επιπέδου. Εξαιρουμένου όμως του πρώτου στρώματος ($l = 1$), για το οποίο, το διάνυσμα εισόδου είναι η διανυσματική τροχιά $\mathbf{V}_{\mathbb{X}_j}$ και όχι ένα προηγούμενο στρώμα :

$$\mathbf{X}_l = \begin{cases} \mathbf{V}_{x_j}, & \text{if } l = 1 \\ L_{l-1}, & \text{if } l > 1 \end{cases}$$

Η αρχιτεκτονική αυτού του τύπου δικτύου περιλαμβάνει δύο στάδια: ένα στάδιο κωδικοποίησης και ένα στάδιο αποκωδικοποίησης. Όσον αφορά το στάδιο της κωδικοποίησης, ο αριθμός των νευρώνων σε κάθε στρώμα μειώνεται, ούτως ώστε η διαστατικότητα του αρχικού διανύσματος εισόδου να μπορεί να μειωθεί αποτελεσματικά όταν περνάει μέσα από το δίκτυο.

Σύμφωνα με την ακόλουθη εξίσωση, το διάνυσμα χαρακτηριστικών που χρησιμοποιείται είναι η έξοδος του στρώματος 2, όταν η τροχιά j χρησιμοποιείται ως διάνυσμα εισόδου στο στρώμα 1:

$$\mathbf{f}_j = L_2$$

Τα εξαγόμενα χαρακτηριστικά \mathbf{f}_j , τροφοδοτούνται σε ένα πλαίσιο συσταδοποίησης που παράγει ένα σύνολο συστάδων \mathbf{C} . Τέλος, για μία εξαγόμενη συστάδα C_n , το αντίστοιχο μοτίβο κίνησης M_n , προσδιορίζεται από μια τροχιά που βασίζεται στην ελαχιστοποίηση της απόστασης από το κεντροειδές της συστάδας.

Στην ίδια κατεύθυνση κινείται και η εργασία των Olive et al. (2020) η οποία παρουσιάζει την δυναμική βαθιάς μάθησης με τη χρήση ενός Autoencoder στην ομαδοποίηση τροχιάς. Απώτερος στόχος της χρήσης Autoencoder είναι η αυτόματη εξαγωγή χαρακτηριστικών με ταυτόχρονη μείωση διαστάσεων μιας και τα χαρακτηριστικά αυτά μπορεί να είναι σε πολύ μικρότερο αριθμό από τις διαστάσεις του αρχικού σετ δεδομένων. Οι συγγραφείς στην εργασία αυτοί εφαρμόζουν για πρώτη φορά αυτή την υβριδική μέθοδο για τον προσδιορισμό και την διαχείριση της εναέριας κυκλοφορίας. Πιο συγκεκριμένα βασίζονται σε ένα σύνολο δεδομένων 19.480 τροχιών που προσγειώνονται στο αεροδρόμιο της Ζυρίχης. Τα αποτελέσματα τους έδειξαν ότι αν και ορισμένα σημεία σχηματίζονται φυσικά στον χώρο μειωμένης διάστασης, υπάρχουν περιπτώσεις όπου τα σημεία δεν διαχωρίζονται καλά και επικαλύπτονται.

Πιο συγκεκριμένα, η εργασία στηρίζεται στους αυτόματους κωδικοποιητές οι οποίοι συνήθως εκπαιδεύονται με οπίσθια διάδοση και χωρίς επίβλεψη. Η συνάρτηση κωδικοποίησης ενός αυτόματου κωδικοποιητή αντιστοιχίζει τα δεδομένα εισόδου $s \in \mathbb{R}^d$ σε μια κρυφή αναπαράσταση $y \in \mathbb{R}^h = e(s)$. Η συνάρτηση αποκωδικοποίησης

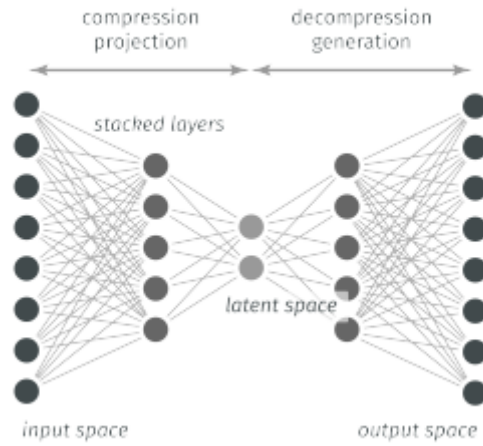
αντιστοιχίζει την κρυφή αναπαράσταση πίσω στον αρχικό χώρο εισόδου σύμφωνα με $\hat{s} = d(y)$. Ο στόχος του μοντέλου αυτόματου κωδικοποιητή είναι να ελαχιστοποιήσει το σφάλμα του ανακατασκευασμένου αποτελέσματος:

$$(w, b, w', b') = \operatorname{argmin} \ell (s, d(e(s)))$$

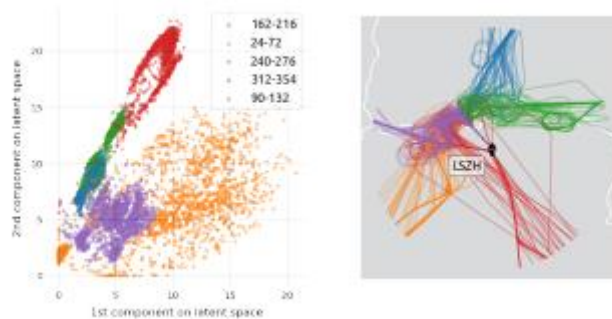
όπου το $\ell(u, v)$ είναι μια συνάρτηση απώλειας που προσδιορίζεται σύμφωνα με το εύρος εισόδου. Συνήθως η απώλεια μέσου τετραγώνου σφάλματος (mean squared error - MSE) δίνεται ως:

$$\ell(u, v) = \frac{1}{n} \sum \|u_i - v_i\|^2 \quad (\text{Εξίσωση 2})$$

Το παρακάτω σχήμα (εικόνα 21) δείχνει ένα παράδειγμα δομής νευρωνικού δικτύου αυτόματης κωδικοποίησης, όπου τα επίπεδα στοιβάζονται, έτσι ώστε να χειρίζονται καλύτερα τον όγκο και την πολυπλοκότητα στα σύνολα δεδομένων (όπως αναλύονται στη μελέτη των συγγραφέων (Olive et al., 2020)). Οι αυτόματοι κωδικοποιητές (που αναφέρονται επίσης με την ευρύτερη έννοια ως μέθοδοι ανακατασκευής) έχουν αποδειχθεί χρήσιμοι τόσο στην ανίχνευση ανωμαλιών σε μεγάλα σύνολα δεδομένων, όσο και σε τροχιές αεροσκαφών. Το σχήμα αυτό απεικονίζει επίσης πώς ολόκληρο το σύνολο δεδομένων προβάλλεται σε έναν δισδιάστατο λανθάνοντα χώρο μέσω του πρώτου τμήματος του νευρωνικού δικτύου αυτόματης κωδικοποίησης. Τα δείγματα προβάλλονται σε έναν χώρο μικρότερης διάστασης, προτού ο χειριστής παραγωγής επιχειρήσει να ανακατασκευάσει τα αρχικά δείγματα. Παρά τη φτωχότερη ανακατασκευαστική ισχύ, οι συγγραφείς (Olive et al., 2020) προέβλεψαν τα δείγματά τους σε έναν δισδιάστατο χώρο για να σχεδιάσουν την κατανομή των προβαλλόμενων δειγμάτων στο Σχήμα 22. Σε αυτό, αντιστοίχισαν χρώματα σε κάθε τροχιά με βάση το έδρανο άξονα και σε σχέση με το αεροδρόμιο της Ζυρίχης, τη στιγμή που ένα αεροσκάφος εισέρχεται σε ακτίνα 40 ναυτικών μιλίων. Κάθε χρώμα σχετίζεται με μια εισερχόμενη ροή στην μικροσυστοιχία ιστών (tissue microarray - TMA).



Σχήμα 21 Δομή ενός νευρωνικού δικτύου αυτόματης κωδικοποίησης. (Πηγή: Olive et al., 2020).



Σχήμα 22 Συσταδοποιημένες τροχιές που προσγειώνονται στο αεροδρόμιο της Ζυρίχης, στον διάδρομο 14, εντός 40 ναυτικών μιλίων από το αεροδρόμιο. (Πηγή: Olive et al., 2020).

Καθώς οι αυτόματοι κωδικοποιητές φαίνεται να διαχωρίζουν με φυσικό τρόπο παρόμοια δεδομένα στον λανθάνοντα χώρο, οι συγγραφείς (Olive et al., 2020) παρουσιάζουν στη συνέχεια δύο μεθόδους για την περαιτέρω επιβολή μιας συσταδοποίησης δεδομένων υψηλών διαστάσεων, με βάση την αναπαράστασή τους σε χώρο χαμηλότερης διάστασης. Οι μέθοδοι αυτοί είναι η artefact (ενσωματωμένη στοχαστική ενσωμάτωση γείτονα, βασισμένη σε αυτόματο κωδικοποιητή για τροχιές αεροσκαφών - AutonecodeR-based t-SNE for Aircraft Trajectories) και η βαθιά συνελκτική ενσωματωμένη συσταδοποίηση (Deep Convolutional Embedded Clustering - DCEC).

Αρχικά, ανέπτυξαν τη μέθοδο artefact που βασίζεται στην t-SNE, ως τρόπο για να επιβάλουν μια πιο προσαρμοσμένη αναπαράσταση στον λανθάνοντα χώρο για τις παραδοσιακές μεθόδους συσταδοποίησης. Δεύτερον, προσάρμοσαν μια υπάρχουσα μέθοδο βαθιάς συσταδοποίησης που ονομάζεται DCEC, (η οποία είχε σχεδιαστεί

αρχικά για εικόνες), ώστε να μπορεί να εφαρμοστεί στην συσταδοποίηση των τροχιών. Ο στόχος είναι να συγκριθούν αυτές οι δύο μέθοδοι όσον αφορά τις δυνατότητες συσταδοποίησης. Το πλεονέκτημα της DCEC έναντι της artefact, είναι ότι μπορεί να αναγνωρίσει τα συμπλέγματα άμεσα, χωρίς την ανάγκη χρήσης μίας κλασικής μεθόδου συσταδοποίησης. Ωστόσο, στη περίπτωση της DCEC, ο αριθμός των συστάδων πρέπει να οριστεί εκ των προτέρων. Εν αντιθέσει με την artefact, όπου αυτός ο ορισμός δεν είναι απαραίτητος εάν η artefact χρησιμοποιείται σε συνδυασμό με κάποιον άλλο αλγόριθμο όπως π.χ. τον αλγόριθμο χωρικής συσταδοποίησης εφαρμογών με θόρυβο με βάση την πυκνότητα (Density-based spatial clustering of applications with noise – DBSCAN). Πιο αναλυτικά, η μέθοδος artefact βασίζεται σε έναν αυτόματο κωδικοποιητή και μια συνάρτηση απώλειας που συνδυάζει το σφάλμα ανακατασκευής και έναν όρο ρύθμισης που βασίζεται στην t-SNE. Η t-SNE είναι μια μέθοδος μείωσης διαστάσεων και επίσης μια ισχυρή τεχνική δισδιάστατης ή τρισδιάστατης απεικόνισης για δεδομένα υψηλών διαστάσεων.

Συσταδοποίηση στον λανθάνοντα χώρο

Η διαδικασία εκπαίδευσης τόσο της artefact όσο και της DCEC συνίσταται στην ελαχιστοποίηση της ακόλουθης συνάρτησης απώλειας που βασίζεται τόσο στην απώλεια δικτύου (σφάλμα ανακατασκευής) όσο και στην απώλεια συσταδοποίησης (απόκλιση Kullback – Leibler - KL):

$$(w, b, w', b') = \operatorname{argmin} \ell(s, d(e(s))) + \lambda d_{KL}(P \parallel Q)$$

Ωστόσο, οι δύο απώλειες συσταδοποίησης είναι στην πραγματικότητα πολύ διαφορετικής φύσεως, όπου καθεμία αποτελεί παράδειγμα των δύο ακόλουθων κατηγοριών που αναφέρονται στη μελέτη των συγγραφέων (Min et al., 2018). Αυτές οι κατηγορίες είναι η βοηθητική απώλεια συσταδοποίησης και η κύρια απώλεια συσταδοποίησης.

Πιο αναλυτικά, η μοναδική χρήση της βοηθητικής απώλειας συσταδοποίησης, είναι να καθοδηγεί το δίκτυο ώστε να μπορέσει να αποκτήσει μια πιο φιλική προς την συσταδοποίηση ενσωμάτωση, από την οποία πρέπει να αναγνωριστούν οι αντίστοιχες συστάδες. Αυτό επιτυγχάνεται εκτελώντας έναν αλγόριθμο συσταδοποίησης μετά τη διαδικασία εκπαίδευσης του δικτύου. Ο όρος ρύθμισης t-SNE που χρησιμοποιείται στη μέθοδο artefact, παίζει ακριβώς αυτόν τον ρόλο, με την πραγματική συσταδοποίηση να εκτελείται στον προσαρμοσμένο λανθάνοντα χώρο από έναν ξεχωριστό αλγόριθμο όπως τον DBSCAN. Από την άλλη πλευρά, μια κύρια απώλεια συσταδοποίησης,

επιτρέπει την άμεση λήψη των αναθέσεων συστάδων μετά τη διαδικασία εκπαίδευσης. Αυτή είναι η περίπτωση της απώλειας σκλήρυνσης από την εκχώρηση συστάδας, η οποία χρησιμοποιείται από την DEC/DCEC. Σε αυτήν, η ανάθεση συστάδας j , για κάθε δείγμα i , μπορεί να ληφθεί απευθείας από την κατανομή Q από το $\text{argmax}_j (q_{ij})$.

Η εργασία των Yao et al., (2018), υλοποιεί ένα υβριδικό μοντέλο ομαδοποίησης τροχιάς το οποίο ενσωματώνει γνώση από μοντέλα βαθιάς μάθησης. Προσπαθεί να λύσει το πρόβλημα των τεχνικών ομαδοποίησης της τροχιάς, οι οποίες συχνά βασίζονται σε μέτρα χωρο-χρονικής ομοιότητας που δεν είναι αμετάβλητα στο χώρο και στο χρόνο. Ως αποτέλεσμα, δεν μπορούν να ανιχνεύσουν συμπλέγματα τροχιάς όπου η ομοιότητα εντός του συμπλέγματος εμφανίζεται σε διαφορετικές περιοχές και χρονικές περιόδους. Επιλύουν το πρόβλημα αυτό με η τεχνική κυλιόμενου παραθύρου (Sliding Window) για να γίνει εξαγωγή ενός συνόλου χαρακτηριστικών δυναμικής συμπεριφοράς που αποτυπώνουν τα αμετάβλητα χαρακτηριστικά του χώρου και του χρόνου των τροχιών. Με τη μονάδα εξαγωγής χαρακτηριστικών, μετατρέπεται κάθε τροχιά σε μια ακολουθία χαρακτηριστικών με σκοπό να γίνει μία περιγραφή σε κινήσεις αντικειμένων και χρησιμοποιείται περαιτέρω ένας αυτόματος κωδικοποιητής του μοντέλου «ακολουθία σε ακολουθία» (sequence to sequence - Seq2Seq), για να γίνουν γνωστές αναπαραστάσεις σε βάθος σταθερού μήκους. Η προτεινόμενη μέθοδος αξιολογείται τόσο σε συνθετικά όσο και σε πραγματικά δεδομένα και παρατηρούνται σημαντικές βελτιώσεις απόδοσης σε σχέση με τις υπάρχουσες μεθόδους. Στην μεθοδολογία αυτή βασίζεται η παρούσα διπλωματική εργασία, ως εκ τούτου περισσότερες λεπτομέρειες παρατίθενται παρακάτω.

Διατύπωση προβλήματος

Δεδομένου ενός συνόλου κινούμενων αντικειμένων $O = \{o_1, o_2, \dots, o_n\}$, κάθε αντικείμενο o_i έχει μια ακολουθία ιστορικού εγγραφών GPS $S_0 = (x_1, x_2, \dots, x_M)$. Εδώ, κάθε στοιχείο x αποτελείται από μια πλειάδα (t_x, l_x, a_x, o_x) , όπου t_x είναι η χρονική σήμανση, l_x είναι ένα διδιάστατο διάνυσμα (γεωγραφικό μήκος και γεωγραφικό πλάτος) που αντιπροσωπεύει τη θέση του αντικειμένου, το a_x είναι ένα σύνολο χαρακτηριστικών που συλλέγονται από άλλους αισθητήρες (π.χ., εάν το αντικείμενο είναι αυτοκίνητο, το a_x μπορεί να περιλαμβάνει την ταχύτητα, την ταχύτητα στροφής και την κατανάλωση καυσίμου) και τέλος το o_x είναι το αναγνωριστικό αντικειμένου. Λαμβάνοντας υπόψη ότι μια ακατέργαστη αλληλουχία

S_0 μπορεί να είναι αραιή στην πράξη, την χωρίζουμε σε ένα σύνολο ακολουθιών τροχιάς $TR_o = (TR_1, TR_2, \dots, TR_n)$ που ορίζεται ως εξής: Έχοντας ως δεδομένο το $S_0 = (x_1, x_2, \dots, x_M)$ και ένα χρονικό διάστημα $\Delta t > 0$, μια επακόλουθη $S_0^T = (x_1, x_{i+1}, \dots, x_{i+k})$ είναι μια τροχιά εάν το S_0^T πληροί τις ακόλουθες προϋποθέσεις: α) $\forall 1 < j \leq k, t_{x_j} - t_{x_{j-1}} \leq \Delta t$ και β) δεν υπάρχει συνέχεια στο S_0 που να περιέχει S_0^T και επίσης να πληροί την προϋπόθεση (α). Κλείνοντας, παρουσιάζεται το πλαίσιο για την εύρεση ομάδων αμετάβλητης τροχιάς χώρου και χρόνου το οποίο είναι μια προσέγγιση χωρίς επίβλεψη αποτελούμενη από τέσσερα επίπεδα:

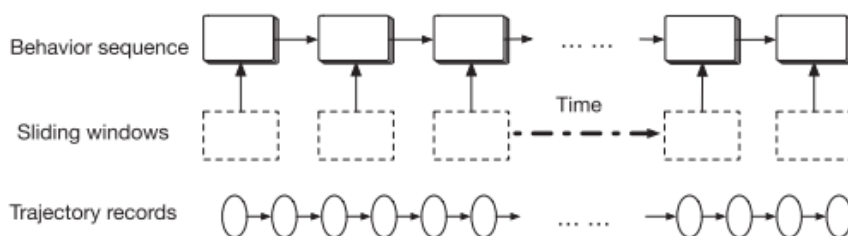
- Επίπεδο προ-επεξεργασίας τροχιάς: Η είσοδος αυτού του επιπέδου είναι οι ακολουθίες εγγραφής GPS των κινούμενων αντικειμένων. Πρέπει ωστόσο να σημειωθεί ότι η ακολουθία είναι θορυβώδης και τα χρονικά διαστήματα μεταξύ ορισμένων ζευγών εγγραφών μπορεί να είναι πολύ μεγάλα. Σε αυτό το επίπεδο, αφαιρούνται οι εγγραφές GPS που είναι χαμηλής ποιότητας και κόβονται οι ακολουθίες σε τροχιές με χρονική συνέχεια.
- Επίπεδο εξαγωγής χαρακτηριστικών κινούμενης συμπεριφοράς: Σε αυτό το επίπεδο, όλες οι τροχιές υποβάλλονται σε επεξεργασία με αλγόριθμο εξαγωγής χαρακτηριστικών κινούμενης συμπεριφοράς. Με βάση τη τεχνική συρόμενου παραθύρου, κάθε τροχιά μετατρέπεται σε μια ακολουθία χαρακτηριστικών.
- Επίπεδο αυτόματου κωδικοποιητή Seq2Seq: Χρησιμοποιείται ένας αυτόματος κωδικοποιητής Seq2Seq για να ενσωματωθεί κάθε ακολουθία χαρακτηριστικών σε ένα διάνυσμα σταθερού μήκους που με τη σειρά του κωδικοποιεί το μοτίβο κίνησης της τροχιάς.
- Στρώμα ανάλυσης συμπλεγμάτων: Τέλος, επιλέγεται ένας κλασικός αλγόριθμος ομαδοποίησης βασισμένος στις πρακτικές ανάγκες και οι ήδη γνωστές αναπαραστάσεις ομαδοποιούνται σε συμπλέγματα.

Διατύπωση προβλήματος

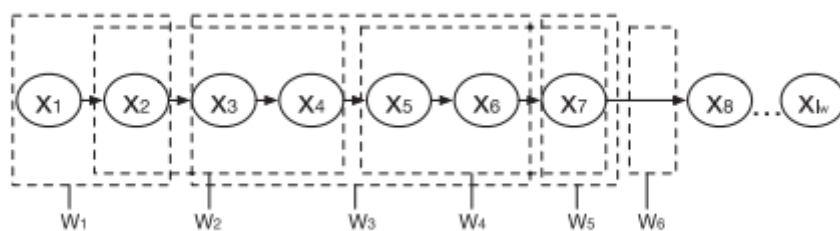
Σε αυτήν την ενότητα, γίνεται επεξεργασία δύο κύριων επιπέδων με βάση το πλαίσιο της έρευνας των συγγραφέων (Yao et al., 2018). Αυτά είναι, το επίπεδο εξαγωγής χαρακτηριστικών και το επίπεδο αυτόματου κωδικοποιητή ακολουθίας σε ακολουθία (sequence to sequence - Seq2Seq). Επιπλέον, γίνεται μια επισκόπηση τυπικών επαναλαμβανόμενων μονάδων που χρησιμοποιείται εξίσου στην έρευνά τους.

Εξαγωγή χαρακτηριστικών με χρήση συρόμενου παραθύρου

Η βασική ιδέα της εξαγωγής χαρακτηριστικών συμπεριφοράς είναι η χρήση της τεχνικής συρόμενου παραθύρου (Sliding Window Technique) για να διασχιστούν οι εγγραφές και να ληφθούν χαρακτηριστικά σε κάθε παράθυρο. Όπως φαίνεται στο Σχήμα 23, με τη βοήθεια της παραπάνω τεχνικής, ο στόχος των συγγραφέων (Yao et al., 2018), είναι να αποκτήσουν χαρακτηριστικά αμετάβλητα του χώρου και του χρόνου για να περιγράψουν τις κινούμενες συμπεριφορές του αντικειμένου. Το L_p και το αντιστάθμισμα (offset) p υποδηλώνουν το πλάτος και τη μετατόπιση του συρόμενου παραθύρου, αντίστοιχα. Πρέπει να σημειωθεί ότι οι κλασικές μέθοδοι συχνά ορίζουν $\text{offset } p = L_p$, αλλά μέσω της έρευνας διαπιστώνεται ότι ένας λεπτότερος βαθμός ανάλυσης του $p = 1/2 \times L_p$ μπορεί να οδηγήσει σε καλύτερη απόδοση. Με αυτόν τον τρόπο, κάθε εγγραφή σε μια τροχιά εκχωρείται σε δύο διαδοχικά παράθυρα και μπορούν να καταγραφούν οι περισσότερες αλλαγές συμπεριφοράς. Επειδή η πυκνότητα εγγραφής είναι ανομοιόμορφη, εισάγονται επίσης ορισμένα εικονικά παράθυρα χωρίς εγγραφές, όπως το W_6 στο σχήμα 24.



Σχήμα 23 - Εξαγωγή κινητικής συμπεριφοράς. (Πηγή: Yao et al., 2018)



Σχήμα 24 - Χρόνος ολίσθησης δημιουργίας παραθύρων. (Πηγή: Yao et al., 2018)

Έπειτα, και σύμφωνα με τη μελέτη των (Yao et al., 2018), γίνεται η περιγραφή της λεπτομερούς διαδικασίας εξαγωγής χαρακτηριστικών σε κάθε συρόμενο παράθυρο ως εξής. Οι μεταβαλλόμενες αλλαγές συμπεριφοράς μπορούν να αντικατοπτρίζονται από

τις διαφορές των χαρακτηριστικών μεταξύ δύο διαδοχικών εγγραφών. Έστω, ότι ένα παράθυρο περιέχει εγγραφές R , συμβολιζόμενες ως $W = (x_1, x_2, \dots, x_R)$ και τα χαρακτηριστικά σε κάθε εγγραφή αποτελούνται από την ταχύτητα και τον ρυθμό στροφής (rate of turn - ROT). Στη συνέχεια, τα εξαγόμενα χαρακτηριστικά των κινούμενων συμπεριφορών περιλαμβάνουν χρονικό διάστημα $\Delta t_i = t_{x_1} - t_{x_{i-1}}$, αλλαγή θέσης $\Delta I_i = I_{x_1} - I_{x_{i-1}}$, αλλαγή ταχύτητας $\Delta s_i = s_{x_1} - s_{x_{i-1}}$ και αλλαγή ROT $\Delta r_i = r_{x_1} - r_{x_{i-1}}$, όπου το i κυμαίνεται από 2 έως R . Με αυτόν τον τρόπο, ένα παράθυρο με εγγραφές R έχει χαρακτηριστικά κινούμενης συμπεριφοράς $R - 1$ ($\Delta I, \Delta s, \Delta r$). Παρόλο που η ταχύτητα και η περιστροφή δεν περιλαμβάνονται ως ρητά χαρακτηριστικά στην ακατέργαστη εγγραφή, μπορούν να υπολογιστούν σύμφωνα με τις πληροφορίες τοποθεσίας. Όπως φαίνεται στο παραπάνω σχήμα, εξετάζεται μια τροχιά με T εγγραφές $TR = (x_1, x_2 \dots x_T)$. Κάθε μία από αυτές έχει μόνο τη χρονική σήμανση και τις συντεταγμένες τοποθεσίας, οι οποίες συμβολίζονται ως $(t, \text{lat}, \text{lon})$. Για την πρώτη εγγραφή x_1 ορίζεται $s_{x_1} = 0$ και $r_{x_1} = 0$. Στη συνέχεια, οι συγγραφείς (Yao et al., 2018) μπορούν να υπολογίσουν τη ταχύτητα και την περιστροφή κάθε εγγραφής κατά

$$s_{x_1} = \frac{\sqrt{(\text{lat}_{x_1} - \text{lat}_{x_{i-1}})^2 + (\text{lon}_{x_1} - \text{lon}_{x_{i-1}})^2}}{t_{x_1} - t_{x_{i-1}}}$$

και

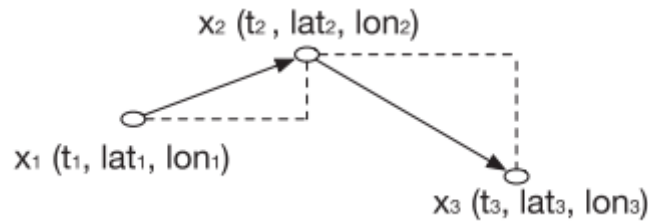
$$r_{x_1} = \arctan \frac{\text{lon}_{x_1} - \text{lon}_{x_{i-1}}}{\text{lat}_{x_1} - \text{lat}_{x_{i-1}}}$$

όπου $i = 2, \dots, T$. Μετά από αυτή τη διαδικασία, η ταχύτητα και η ROT μπορούν να προκύψουν για κάθε τροχιά. Αν $R \geq 1$, για κάθε i από 1 έως R , υπολογίζεται $\Delta t_i, \Delta I_i, \Delta s_i$, και Δr_i . Υπολογίζουν περαιτέρω τον ρυθμό μεταβολής αυτών των χαρακτηριστικών $f_1 = (f_{\Delta I}, f_{\Delta s}, f_{\Delta r})$, όπου $f_{\Delta I} = \Delta I_i / \Delta t_i$, $f_{\Delta s} = \Delta s_i$ και $f_{\Delta r} = \Delta r_i$. Για δύο διαδοχικές εγγραφές, το $f_{\Delta I}$ δηλώνει τη μέση ταχύτητα, το $f_{\Delta s}$ δηλώνει την αλλαγή των ταχυτήτων και το $f_{\Delta r}$ αντιπροσωπεύει τη μεταβολή των ROT. Αφού υπολογίσουν αυτές τις δυνατότητες σε κάθε ζεύγος, δημιουργούν ένα σύνολο χαρακτηριστικών $f = \{f_1, f_2, \dots, f_R\}$. Έπειτα, χρησιμοποιούν τα στατιστικά στοιχεία για να δημιουργήσουν τις δυνατότητες στο συρόμενο παράθυρο. Εδώ, επιλέγονται έξι στατιστικά στοιχεία {μέση, μέγιστη, 75% ποσοτική, 50% ποσοτική, 25% ποσοτική,

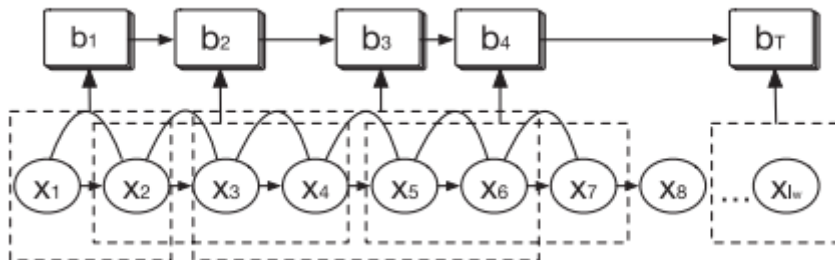
ελάχιστη}. Συνοψίζοντας, με βάση τα χαρακτηριστικά κινούμενης συμπεριφοράς του, κάθε παράθυρο b έχει διαστάσεις $3 \times 6 = 18$, αποτελούμενο από

$$\{f_{\Delta t}, f_{\Delta s}, f_{\Delta r}\} \times \{\text{mean}, \text{max}, 75\% \text{quantile}, 50\% \text{quantile}, 25\% \text{quantile}, \text{min}\}$$

Εάν $R = 0$, παραλείπουν αυτό το παράθυρο. Το σχήμα 26 και ο αλγόριθμος 1 δείχνουν τη διαδικασία δημιουργίας ακολουθίας χαρακτηριστικών κινούμενης συμπεριφοράς. Για κάθε τροχιά στο T , δημιουργούν πρώτα την ακολουθία κινούμενης συμπεριφοράς. Στη συνέχεια, βάζουν αυτές τις ακολουθίες σε ένα σύνολο και το δηλώνουν ως $S = \{B_{TR_1}, B_{TR_2}, \dots, B_{TR_N}\}$. Τέλος, ομαλοποιούν κάθε δυνατότητα για να προετοιμαστούν για το επόμενο επίπεδο αυτόματης κωδικοποίησης ακολουθίας σε ακολουθία.



Σχήμα 25 - Πλήρης γνωρίσματα. (Πηγή: Yao et al., 2018)



Σχήμα 26 - Η δημιουργία ακολουθίας κινούμενης συμπεριφοράς. (Πηγή: Yao et al., 2018)

Χρήση των Deep Neural Network μεθόδων

Τα επαναλαμβανόμενα νευρωνικά δίκτυα (Recurrent neural network – RNN) είναι κατάλληλα για τη μοντελοποίηση αλληλουχιών εισόδου με μεταβλητό μήκος, των οποίων οι επαναλαμβανόμενες μονάδες σχηματίζουν έναν κατευθυνόμενο κύκλο. Για μια ακολουθία εισόδου $X = (x_1, x_2, \dots, x_T)$ όπου $i \in [1, N]$, τα RNN ενημερώνουν την κρυφή κατάσταση h_t σύμφωνα με την τρέχουσα είσοδο x_t και την προηγούμενη κρυφή κατάσταση h_{t-1} . Η κρυφή κατάσταση h_t λειτουργεί ως εσωτερική μνήμη τη στιγμή

που συλλαμβάνει δυναμικές χρονικές πληροφορίες στην ακολουθία. Συγκεκριμένα, η κρυφή κατάσταση ενημερώνεται ως

$$h_t = f(h_{t-1}, x_t)$$

όπου $f(\cdot)$ είναι η συνάρτηση ενεργοποίησης. Πρέπει να σημειωθεί ότι η συνάρτηση $f(\cdot)$ χρησιμοποιείται για τον υπολογισμό ενός σταθμισμένου αθροίσματος εισόδων και την εφαρμογή ενός μη γραμμικού μετασχηματισμού, ο οποίος εξαρτάται από τον τύπο των επαναλαμβανόμενων μονάδων. Καθώς το μη τροποποιημένο RNN δυσκολεύεται στη πράξη, κατά την εκμάθηση μακροπρόθεσμων εξαρτήσεων, οι συγγραφείς (Yao et al., 2018) χρησιμοποιούν τόσο το δίκτυο μακράς βραχύχρονης μνήμης (Long short-term memory – LSTM) όσο και τον μηχανισμό κλειστής επαναλαμβανόμενης μονάδας (gated recurrent unit – GRU) για να ξεπεράσουν αυτή την έλλειψη. Παρακάτω περιγράφονται εν συντομία τα προγράμματα ενημέρωσης στα LSTM και GRU.

LSTM Δίκτυο

Δεδομένης μιας ακολουθίας εισόδου $X = (x_1, x_2, \dots, x_T)$, το LSTM υπολογίζει την κρυφή διανυσματική ακολουθία $H = (h_1, h_2, \dots, h_T)$ διατηρώντας μια μνήμη h_{t-1} τη χρονική στιγμή t και αποφασίζοντας αν θα διατηρήσει την υπάρχουσα μνήμη,

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ \bar{c}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ c_t &= f_t \cdot c_{t-1} + i_t \cdot \bar{c}_t \\ h_t &= o_t \cdot \tanh(c_t) \end{aligned}$$

όπου $\sigma(\cdot)$ είναι η σιγμοειδής συνάρτηση, c_t είναι το περιεχόμενο μνήμης της μονάδας και το \bar{c}_t δηλώνει το νέο περιεχόμενο μνήμης. Υπάρχουν τρεις πύλες στο LSTM: η πύλη εισόδου i_t , η πύλη μνήμης f_t και η πύλη εξόδου o_t . Σημειώνεται ότι, η πύλη εισόδου i_t διαμορφώνει τον βαθμό στον οποίο πρέπει να αποθηκευτεί η νέα μνήμη, η πύλη μνήμης f_t ρυθμίζει τον βαθμό στον οποίο πρέπει να ξεχαστεί η υπάρχουσα μνήμη και η πύλη εξόδου o_t διαμορφώνει την ποσότητα περιεχομένου μνήμης για έκθεση. Διαισθητικά, το LSTM ανιχνεύει το σημαντικό μέρος της εισόδου και το αποθηκεύει για μεγάλο χρονικό διάστημα. Ως εκ τούτου, είναι κατάλληλο για εκμάθηση μακροχρόνιας εξάρτησης, η οποία έχει αποδειχθεί επιτυχής σε πολλές εφαρμογές όπως

προκύπτει για παράδειγμα, μέσα από τις μελέτες των (Chung et al., 2016) και (Dai & Le, 2015).

GRU Δίκτυο

Ο μηχανισμός νευρωνικού δικτύου GRU, υιοθετείται ευρέως στην αυτόματη μετάφραση που βασίζεται σε κωδικοποιητή και αποκωδικοποιητή. Επιτρέπει σε κάθε επαναλαμβανόμενη μονάδα να καταγράφει προσαρμοστικά τις εξαρτήσεις διαφορετικών χρονικών κλιμάκων. Όντας διαφορετικός από το LSTM, ο GRU ελέγχει τη ροή πληροφοριών μέσα στη μονάδα χωρίς ξεχωριστό κελί μνήμης. Πιο επίσημα

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\bar{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t])$$

$$h_t = (1 - z_t)h_{t-1} + z_t\bar{h}_t$$

Ο GRU έχει μόνο δύο πύλες: την πύλη ενημέρωσης z_t και την πύλη επαναφοράς r_t . Η κρυφή κατάσταση h_t είναι μια γραμμική παρεμβολή των h_{t-1} και \bar{h}_t . Ως συνδυασμός πύλης εισόδου και πύλης μνήμης στο LSTM, η πύλη ενημέρωσης z_t , ρυθμίζει κατά πόσο αποθηκεύει ή ξεχνάει νέες πληροφορίες η αντίστοιχη μονάδα. Με αυτόν τον τρόπο, ο GRU συγχωνεύει το περιεχόμενο μνήμης c στην κρυφή κατάσταση και επιτυγχάνει έναν απλούστερο αρχιτεκτονικό σχεδιασμό.

Αυτόματος κωδικοποιητής Seq2Seq

Σε αυτήν την ενότητα και με βάση τη μελέτη τους, οι συγγραφείς (Yao et al., 2018) περιγράφουν ένα μοντέλο που χρησιμοποιεί αυτόματο κωδικοποιητή ακολουθίας σε ακολουθία για την ανακατασκευή της ακολουθίας κινούμενης συμπεριφοράς και δημιουργεί μια βαθιά αναπαράσταση της τροχιάς σταθερού μήκους. Το μοντέλο αυτόματης κωδικοποίησης Seq2Seq αποτελείται από δύο RNN: τον κωδικοποιητή RNN, ο οποίος απεικονίζεται στο αριστερό μέρος του σχήματος 27 και τον αποκωδικοποιητή RNN που απεικονίζεται στη δεξιά πλευρά. Η είσοδος του μοντέλου είναι μια ακολουθία συμπεριφοράς $B_{TR_1} = (b_1, b_2, \dots, b_T)$. Ο κωδικοποιητής RNN διαβάξει διαδοχικά την ακολουθία εισόδου και ενημερώνει αναλόγως την κρυφή κατάσταση h_t . Ο κωδικοποιητής RNN ενημερώνεται ως

$$h_t = f(h_{t-1}, b_t)$$

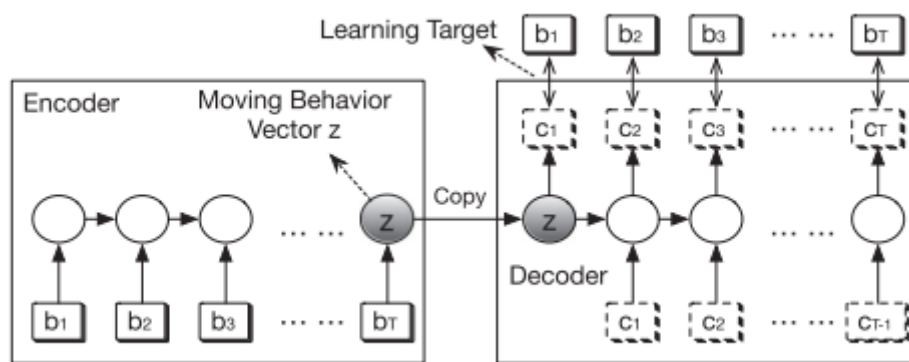
Μετά την επεξεργασία του τελευταίου b_T , η κρυφή κατάσταση h_T χρησιμοποιείται ως αναπαράσταση για ολόκληρη την ακολουθία. Στη συνέχεια, ο αποκωδικοποιητής δημιουργεί πρώτα την έξοδο c_1 λαμβάνοντας το h_T ως αρχικοποιημένη κρυφή κατάσταση του αποκωδικοποιητή RNN και παράγει περαιτέρω (c_2, c_3, \dots, c_T) . Ο αποκωδικοποιητής RNN ενημερώνεται ως

$$h^d_t = f(h^d_{t-1}, c_{t-1}, h_T)$$

Ο στόχος του αποκωδικοποιητή είναι η ανακατασκευή της ακολουθίας εισόδου $B_{TR_1} = (b_1, b_2, \dots, b_T)$. Με άλλα λόγια, ο κωδικοποιητής RNN και ο αποκωδικοποιητής RNN εκπαιδεύονται μαζί ελαχιστοποιώντας το σφάλμα ανασυγκρότησης, μετρούμενο με το γενικό μέσο τετραγωνικό σφάλμα:

$$\text{MSE} = \sum_{t=1}^T \|b_t - c_t\|^2$$

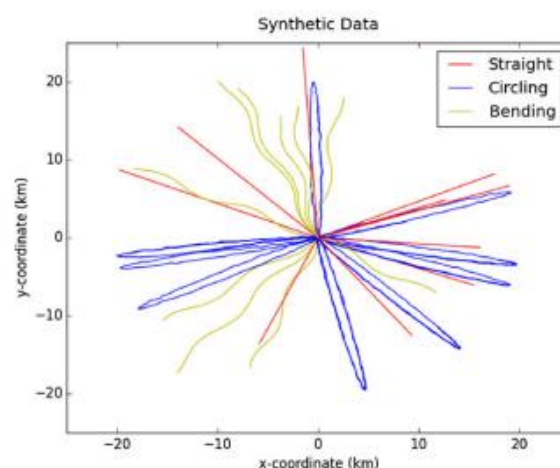
Καθώς η ακολουθία εισόδου λαμβάνεται ως μαθησιακός στόχος, η διαδικασία κατάρτισης δεν χρειάζεται δεδομένα με ετικέτα. Το διάνυσμα συμπεριφοράς κινούμενου σταθερού μήκους z είναι μια σημαντική αναπαράσταση για την ακολουθία συμπεριφοράς εισόδου B_{TK} επειδή η ακολουθία εισόδου οπής μπορεί να αναδημιουργηθεί από το z με τον αποκωδικοποιητή. Μετά από αυτή τη διαδικασία, οι (Yao et al., 2018) παίρνουν το σύνολο διανυσμάτων κινούμενης συμπεριφοράς $Z = \{z_{TR_1}, z_{TR_2}, \dots, z_{TR_N}\}$. Στη συνέχεια, το τροφοδοτούν με έναν κλασικό αλγόριθμο ομαδοποίησης, όπως τα K-μέσα (K-means) και τέλος λαμβάνουν τα συμπλέγματα.



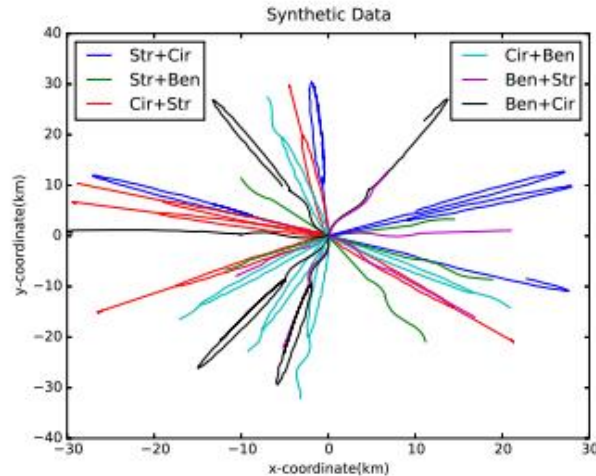
Σχήμα 27 - Αρχιτεκτονική του αυτόματου κωδικοποιητή Seq2Seq. (Πηγή: Yao et al., 2018)

Η μεθοδολογία αυτή αξιολογήθηκε σε συνθετικά και πραγματικά σύνολα δεδομένων για να ελέγξουν την αποτελεσματικότητά του. Για το σύνολο συνθετικών δεδομένων, προσομοίωσαν 9.000 τροχιές που περιλαμβάνουν εννέα είδη μοτίβων κίνησης που

αποτελούνται από τρία είδη βασικών μοτίβων κίνησης {Ευθεία, Κύκλος, Κάμψη} και έξι είδη μοτίβων συνδυασμού {Ευθεία + Κύκλος, Ευθεία + Κάμψη, Κύκλος + Κάμψη, Κύκλος + Ευθεία, Κάμψη + Ευθεία, Κάμψη + Κύκλος}. Κάθε μοτίβο έχει 1.000 τροχιές. Η συχνότητα δειγματοληψίας και το χρονικό μήκος κάθε τροχιάς δημιουργήθηκαν τυχαία από 2.500 έως 5.000 δευτερόλεπτα. Αφού δημιούργησαν τις τροχιές, υπολόγισαν τα χαρακτηριστικά της θέσης με τις εξισώσεις 1 και 2. Επιπλέον, πρόσθεσαν θόρυβο Gaussian στη διαδικασία δημιουργίας τοποθεσίας. Μέρος του συνόλου συνθετικών δεδομένων παρουσιάζεται στα σχήματα 28 και 29. Το πραγματικό σύνολο δεδομένων αντιστοιχεί σε 200 σκάφη στην Κίνα, που περιέχουν 50 φορτηγά σκάφη, 50 αλιευτικά σκάφη, 50 σκάφη πετρελαίου και 50 επιβατηγά σκάφη. Τα δεδομένα κίνησης του σκάφους συλλέγονται από το Σύστημα Αυτόματης Ταυτοποίησης (Automatic Identification System - AIS). Το AIS είναι ένας από τους σημαντικότερους τρόπους ευαισθητοποίησης στον τομέα της ναυτιλίας. Τα μηνύματα AIS μπορούν να χωριστούν σε δυναμικά μηνύματα και στατικά. Τα δυναμικά μηνύματα αναφέρουν τη δυναμική κατάσταση του σκάφους, συμπεριλαμβανομένου του χρόνου, της θέσης (γεωγραφικό μήκος και γεωγραφικό πλάτος), την πορεία πάνω από το έδαφος, την ταχύτητα πάνω από το έδαφος και την κατεύθυνση. Τα στατικά μηνύματα περιλαμβάνουν τον τύπο, το όνομα και το μέγεθος. Ο χρόνος εγγραφής είναι από τον Μάιο του 2016 έως τον Ιούνιο του 2016. Υπάρχουν συνολικά 5.924.142 εγγραφές στο σύνολο δεδομένων. Μετά τον διαχωρισμό της τροχιάς, οι (Yao et al., 2018) δημιούργησαν 4.700 τροχιές.



Σχήμα 28 - Μέρος των συνθετικών βασικών μοτίβων κίνησης που αποτελούνται από 10 ευθείες τροχιές, 10 κυκλικές τροχιές και 10 καμπύλες. (Πηγή: Yao et al., 2018)



Σχήμα 29 - Μέρος του συνθετικού συνδυασμού μοτίβων κίνησης. Κάθε μοτίβο συνδυασμού έχει πέντε τροχιές. Str = Ευθεία, Cir = Κύκλος, Ben = Κάμψη. (Πηγή: Yao et al., 2018)

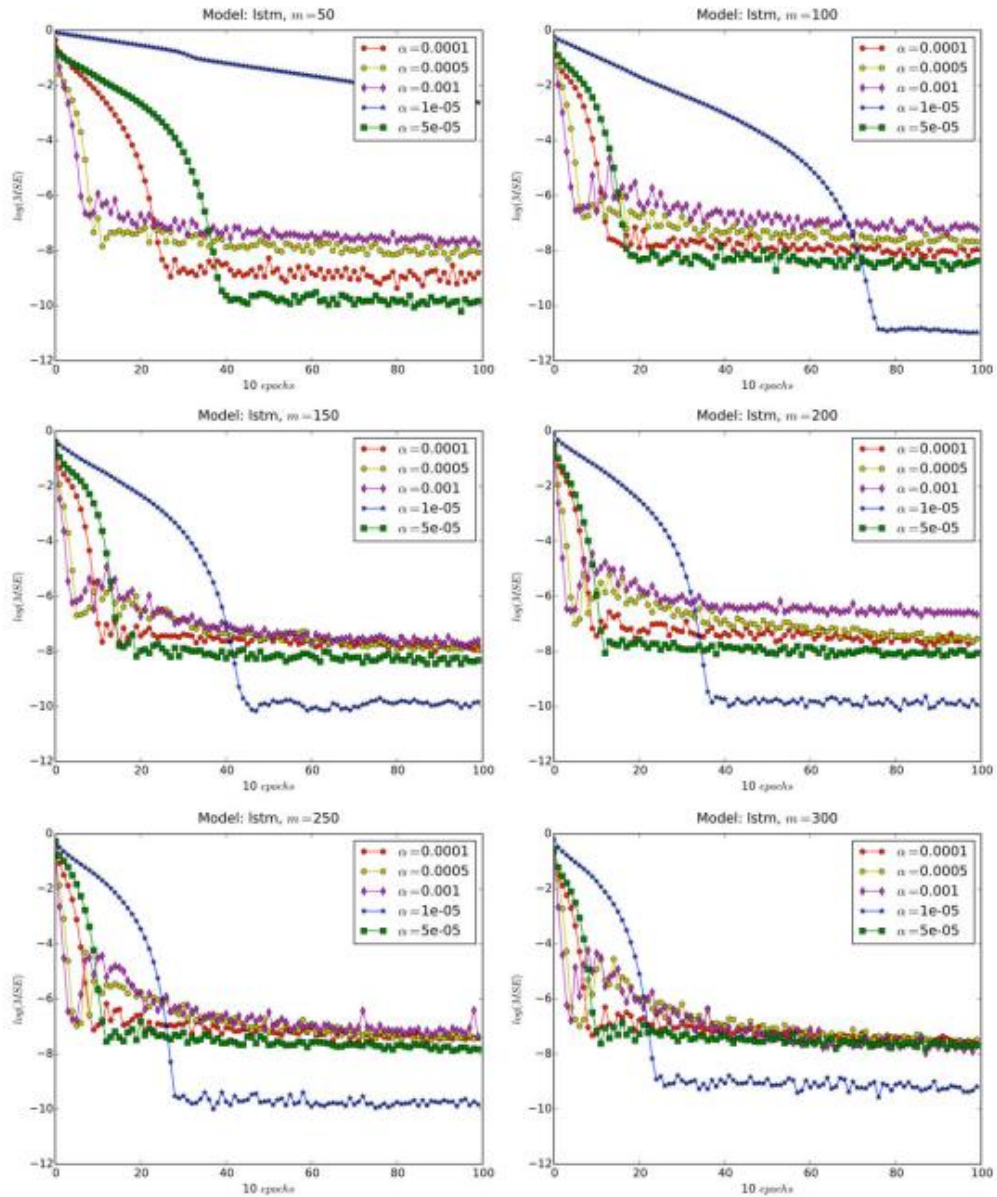
Για τη ρύθμιση παραμέτρων, τόσο το LSTM όσο και ο GRU έχουν τρεις κύριες παραμέτρους: (α) το ρυθμό εκμάθησης α που ελέγχει το μέγεθος του βήματος ενημέρωσης της παραμέτρου, (β) το μέγεθος της κρυφής κατάστασης m , το οποίο ελέγχει το μέγεθος της ενσωμάτωσης της τροχιάς και (γ) τον αριθμό των εποχών κατάρτισης n . Επιτρέποντας την επίδραση διαφορετικών παραμέτρων, οι συγγραφείς (Yao et al., 2018) ορίζουν τις προεπιλεγμένες τιμές των παραμέτρων ως εξής: (α) $\alpha = 0.00001$, $m = 250$ και $n = 400$ για LSTM. και (β) $\alpha = 0.00001$, $m = 100$, και $n = 800$ για GRU. Για το επίπεδο εξαγωγής συμπεριφοράς, ορίζουν το συρόμενο παράθυρο ως 600 s και το αντιστάθμισμα του παραθύρου ως 300 s. Εφάρμοσαν το πλαίσιο με Python και TensorFlow. Όλα τα πειράματα πραγματοποιήθηκαν σε διακομιστή με Intel Xeon CPU 2,10 GHz. Τα δεδομένα και ο κώδικας είναι διαθέσιμα δημόσια. Στη συνέχεια, συγκρίνουν τη μέθοδό τους, συμπεριλαμβανομένων των LSTM και GRU, με τέσσερις μεθόδους ομαδοποίησης τροχιάς που βασίζονται σε διαφορετικά μέτρα, συμπεριλαμβανομένων της μεγαλύτερης κοινής υπακολουθίας (Longest Common Subsequence – LCSS), της δυναμικής στρέβλωσης χρόνου (Dynamic Time Warping – DTW), της ανταπόκρισης έκτακτης ανάγκης (emergency demand response – EDR) και την απόσταση Hausdorff (Hausdorff distance). Όλες οι συναρτήσεις απόστασης μπορούν να χειριστούν τροχιές με διαφορετικά μήκη. Σύμφωνα με την μελέτη των (Besse et al., 2016), οι LCSS, DTW και EDR είναι λειτουργίες απόστασης που βασίζονται σε στρέβλωση και στοχεύουν στην επίλυση του προβλήματος μετατόπισης χρόνου. Είναι σε θέση να αντιστοιχίσουν τοποθεσίες από διαφορετικές τροχιές με διαφορετικά ευρετήρια. Αντίθετα, η απόσταση Hausdorff είναι η απόσταση που

βασίζεται σε σχήμα. Για κάθε μέτρο, οι συγγραφείς (Yao et al., 2018), επιλέγουν τον K-Medoids ως αλγόριθμο ομαδοποίησης. Για τα συνθετικά δεδομένα, επειδή είναι γνωστός ο αριθμός των κινούμενων προτύπων συμπεριφοράς, ορίζουν τον αριθμό των συμπλεγμάτων ως 9. Ενώ για τα δεδομένα της πραγματικής ζωής, συντονίζουν τον αριθμό των συμπλεγμάτων και αναλύουν τα αντίστοιχα αποτελέσματα για να επιλέξουν το καλύτερο. Μετρούν τα αποτελέσματα του συμπλέγματος σε ακρίβεια, ανάκληση και ευστοχία. Για κάθε μέθοδο, υπολογίζουν πρώτα την καλύτερη αντιστοιχία μεταξύ των αποτελεσμάτων ομαδοποίησης και των μοτίβων κίνησης εδάφους-αλήθειας. Στη συνέχεια, για κάθε μοτίβο κίνησης, μετρούν την ακρίβεια και την ανάκληση. Η ακρίβεια p και η ανάκληση r υπολογίζονται ως

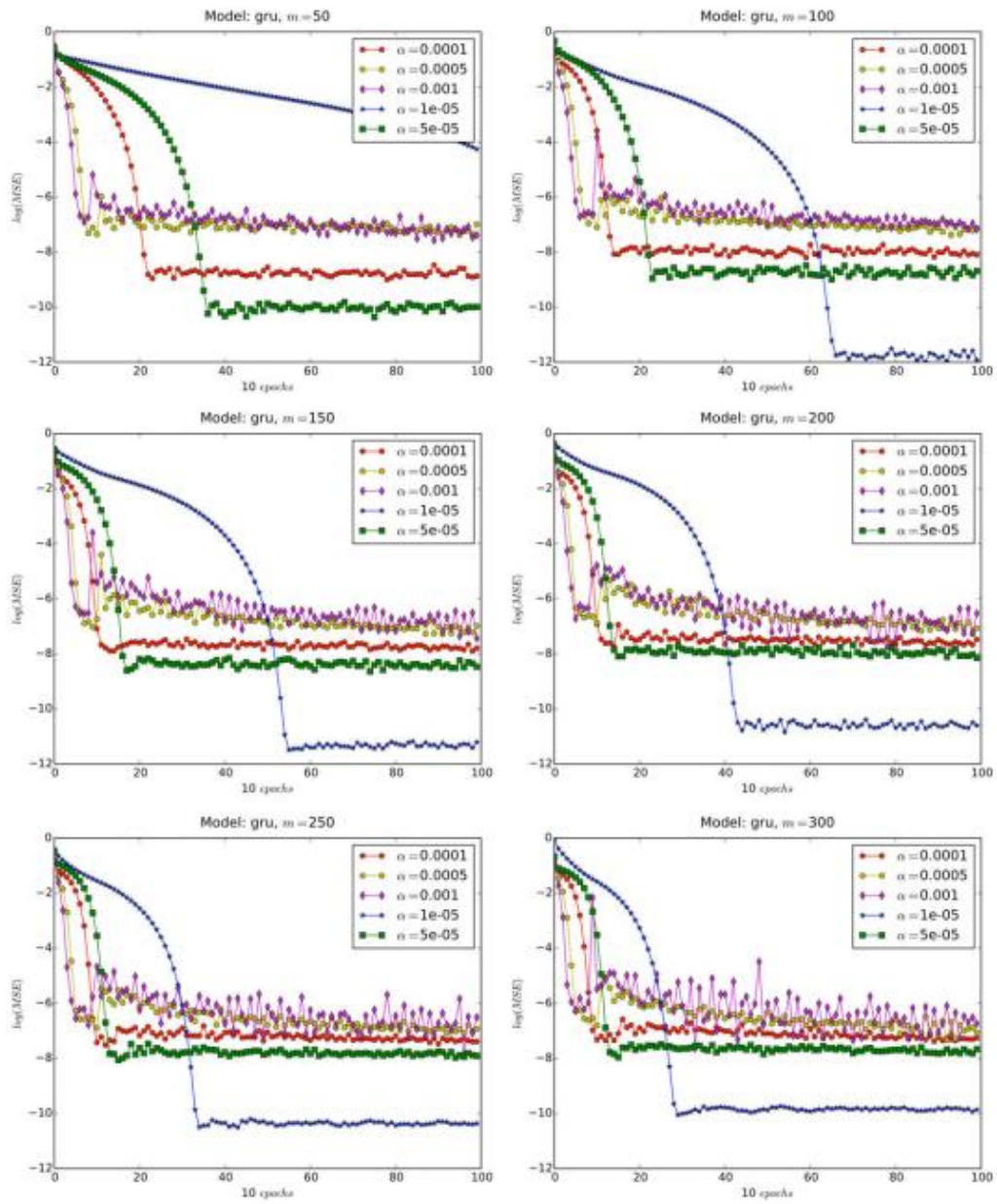
$$p = \frac{TP}{TP + FP}, r = \frac{TP}{TP + FN} \quad (\text{Εξίσωση 17})$$

Εδώ, η πραγματική θετική (true positive - TP) αντιπροσωπεύει τον αριθμό των τροχιών που ταιριάζουν με το μοτίβο κίνησης. Τέλος, μετρούν την ακρίβεια κάθε μεθόδου, που υπολογίζεται ως εξής: Ακρίβεια = Άθροισμα όλων των TP / Αριθμός τροχιών.

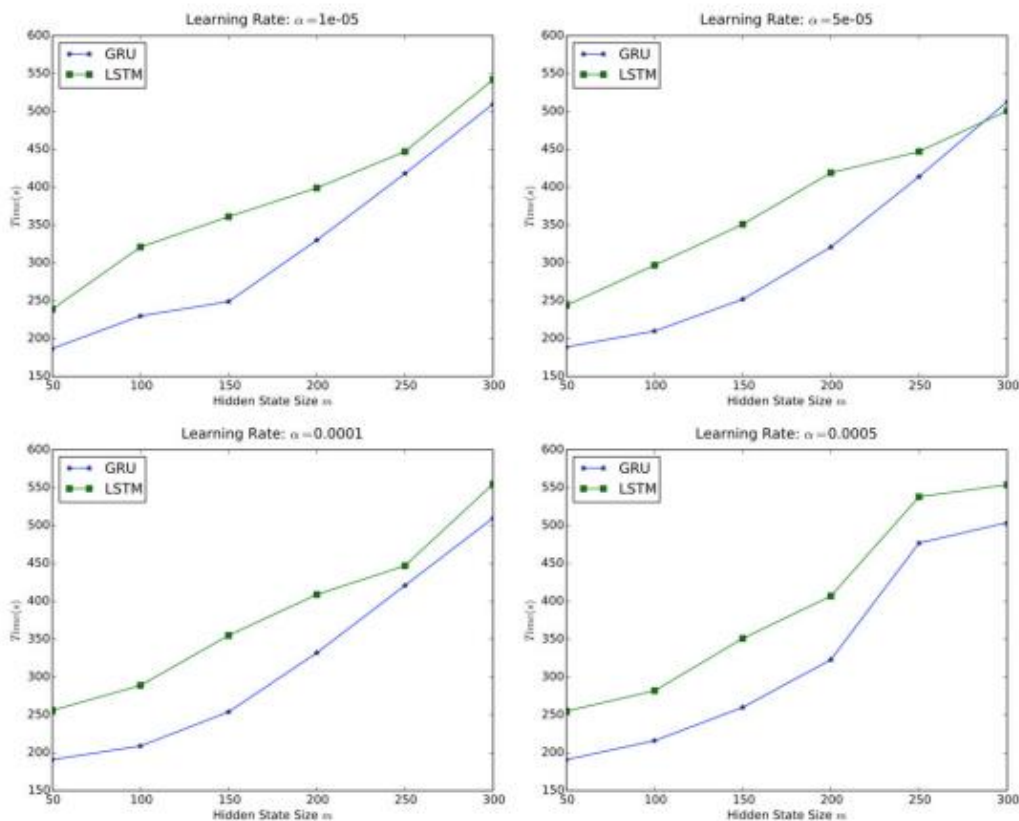
Οι (Yao et al., 2018) μελετούν τις επιδράσεις διαφορετικών παραμέτρων: ο ρυθμός μάθησης α , το μέγεθος κρυφής κατάστασης m και ο αριθμός των εποχών εκπαίδευσης n . Συντονίζουν τις παραμέτρους ως εξής: $\alpha = [0.00001, 0.00005, 0.0001, 0.0005]$, $m = [50, 100, 150, 200, 250, 300]$ και $n = 1,000$. Κατά τη μελέτη της επίδρασης μιας παραμέτρου, διορθώνουν τις άλλες παραμέτρους στις προεπιλεγμένες τιμές τους. Για διαφορετικές ρυθμίσεις, μετρούν το σφάλμα κατάρτισης στο μέσο τετράγωνο σφάλμα (mean squared error – MSE). Έπειτα, επιλέγουν τυχαία 40 τροχιές τόσο από συνθετικά όσο και από δεδομένα πραγματικής ζωής και μετρούν το μέσο σφάλμα εκπαίδευσης σε κάθε εποχή. Τα αποτελέσματα φαίνονται στα σχήματα 30, 31 και 32. Γενικά, το σφάλμα εκπαίδευσης αρχικά μειώνεται και στη συνέχεια γίνεται σταθερό ενώ τα m και n αυξάνονται. Το ποσοστό μάθησης α παίζει επίσης σημαντικό ρόλο. Για να επιτύχουν μια αντιστάθμιση μεταξύ αποτελεσματικότητας και αποδοτικότητας, ορίσαν $m = 250$, $n = 400$ και $\alpha = 0.00001$ για LSTM και $m = 100$, $n = 800$ και $\alpha = 0.00001$ για GRU.



Σχήμα 30 - Το αρχείο καταγραφής μακροπρόθεσμης μνήμης (LSTM) αλλάζει με διαφορετικές παραμέτρους (MSE) .
 (Πηγή: Yao et al., 2018)



Σχήμα 31 - Το αρχείο καταγραφής επαναλαμβανόμενων μονάδων (GRU) αλλάζει με διαφορετικές παραμέτρους (MSE). (Πηγή: Yao et al., 2018)



Σχήμα 32 - Χρονοβόρες αλλαγές με διαφορετικές παραμέτρους. (Πηγή: Yao et al., 2018)

Για τα συνθετικά δεδομένα, οι (Yao et al., 2018) ορίζουν το συρόμενο παράθυρο στα 600 s και το αντιστάθμισμα του παραθύρου στα 300 s. Στη συνέχεια επιλέγουν ξανά τον αλγόριθμο K-means για να δημιουργήσουν τα συμπλέγματα τροχιάς. Το EDR και το LCSS χρειάζονται ένα κατώτατο όριο απόστασης για να καθορίσουν εάν ταιριάζουν δύο εγγραφές. Μετά τον συντονισμό, έθεσαν το όριο στα 100 m. Η απόδοση ομαδοποίησης διαφορετικών μεθόδων φαίνεται στον Πίνακα 1. Τα αποτελέσματα δείχνουν ότι η μέθοδος τους μπορεί να εξαγάγει μοτίβα κίνησης πολύ καλύτερα από τα EDR, LCSS, Hausdorff και DTW. Χρησιμοποιώντας την προσέγγισή τους, οι τροχιές με παρόμοιες κινούμενες συμπεριφορές ομαδοποιούνται, ακόμη και αν η ομοιότητα συμβαίνει σε διαφορετικές περιοχές και χρονικές περιόδους. Εν κατακλείδι, η μέθοδος των συγγραφέων (Yao et al., 2018), έχει βελτιώσει την ακρίβεια κατά περισσότερο από 10% σε σχέση με τις άλλες μεθόδους που χρησιμοποιούνται για συγκρίσεις.

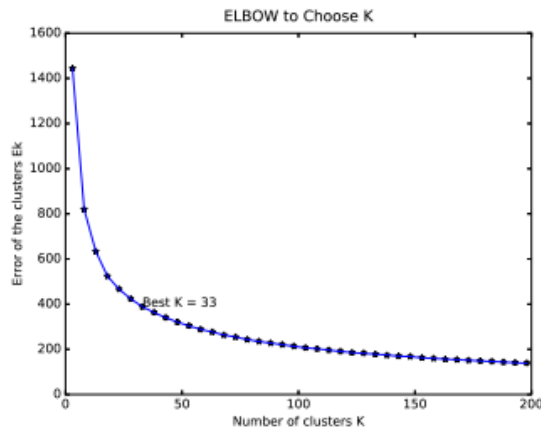
Πίνακας 1 - Ομαδοποίηση της απόδοσης σε συνθετικά δεδομένα. (Πηγή: Yao et al., 2018)

	EDR	LCSS	DTW	Hausdorff	GRU-s2s	LSTM-s2s
Straight	0.465/0.563	0.460/0.411	0.411/0.613	0.423/0.263	0.643/0.723	0.760/0.703
Circling	0.550/0.482	0.610/0.643	0.540/0.462	0.415/0.531	0.768/0.756	0.766/0.823
Bending	0.668/0.678	0.621/0.392	0.472/0.322	0.465/0.379	0.733/0.546	0.652/0.752
Straight+Circling	0.359/0.468	0.573/0.523	0.503/0.474	0.507/0.414	0.571/0.684	0.596/0.410
Straight+Bending	0.453/0.427	0.462/0.574	0.507/0.746	0.435/0.510	0.646/0.823	0.783/0.763
Circling+Bending	0.600/0.581	0.469/0.313	0.766/0.480	0.389/0.283	0.563/0.522	0.738/0.685
Circling+Straight	0.470/0.434	0.388/0.661	0.595/0.377	0.348/0.429	0.664/0.312	0.621/0.891
Bending+Straight	0.327/0.374	0.409/0.582	0.769/0.379	0.375/0.534	0.609/0.927	0.500/0.316
Bending+Circling	0.674/0.419	0.528/0.251	0.525/0.879	0.442/0.387	0.715/0.539	0.688/0.819
Overall accuracy (%)	49.18	48.33	53.69	41.64	64.80	68.47

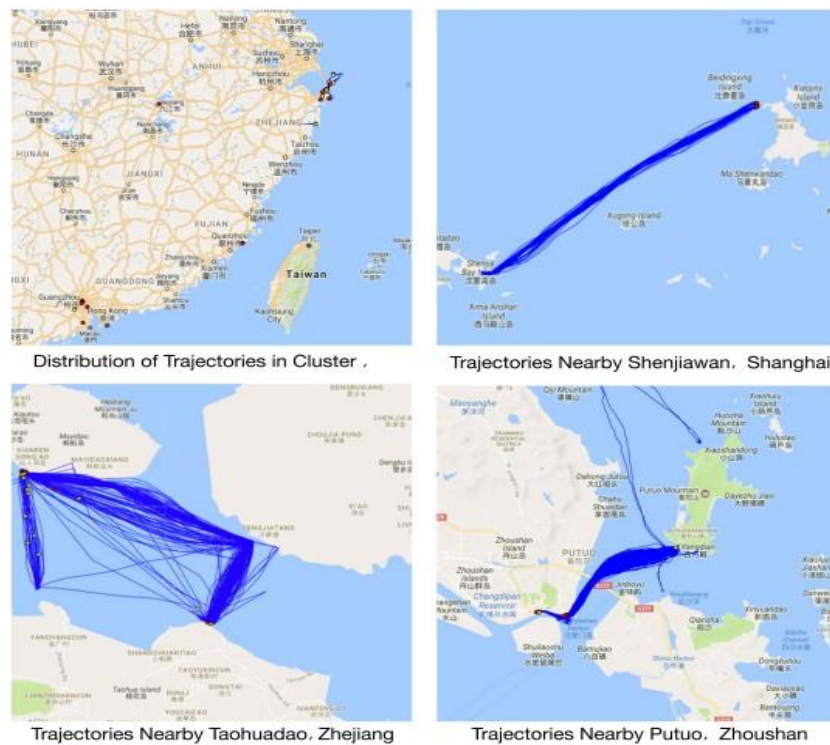
Οι συγγραφείς (Yao et al., 2018), εκτελούν δύο εργασίες σε αυτό το σύνολο δεδομένων. Η πρώτη είναι η τυπική εργασία ομαδοποίησης τροχιάς, όπου χρησιμοποιούν το πλαίσιο τους για να δημιουργήσουν ομάδες που έχουν παρόμοια κινούμενη συμπεριφορά και στη συνέχεια αναλύουν το νόημα των τροχιών σε αυτές. Η δεύτερη είναι η ανάλυση τύπου σκάφους στην οποία εξετάζουν αν τα σκάφη που κατέχουν τον ίδιο τύπο ομαδοποιούνται στο ίδιο σύμπλεγμα ή όχι και έπειτα μετρούν τις ακρίβειες. Χρησιμοποιώντας το πλαίσιο τους, δημιουργούνται διανύσματα συμπεριφοράς κινούμενης τροχιάς. Οι παράμετροι που χρησιμοποιούνται σε αυτήν τη διαδικασία είναι ίδιες με το σύνολο συνθετικών δεδομένων. Παρατηρώντας ότι το LSTM είναι καλύτερο από τον GRU, επιλέγουν το μοντέλο LSTM σε αυτήν την εργασία τους. Επίσης, χρησιμοποιούν τον K-means για να δημιουργήσουν τα συμπλέγματα για το σύνολο Z . Καθώς ο αριθμός των συμπλεγμάτων αληθινής βάσης στα πραγματικά δεδομένα είναι άγνωστος, η διαδικασία επιλογής της τιμής του K διεξάγεται ως εξής. Ρυθμίζουν την τιμή του K από 3 έως 100. Για κάθε K , υπολογίζουν το άθροισμα των αποστάσεων από τα δείγματα στο πλησιέστερο κέντρο βάρους τους και το σημειώνουν ως E_k . Το αποτέλεσμα φαίνεται στο σχήμα 33 που απεικονίζει ότι η τιμή K αντιστοιχεί στο σημείο του αγκώνα (elbow point), το οποίο εντοπίζεται μέσω της μεθόδου αγκώνα (elbow method) που χρησιμοποιείται για τον προσδιορισμό του αριθμού των συμπλεγμάτων σε ένα σύνολο δεδομένων. Όπως φαίνεται στο σχήμα αυτό, επιλέγουν $K = 33$, εξάγοντας 33 συμπλέγματα για τις 4.700 τροχιές. Μερικά από τα αποτελέσματα των συμπλεγμάτων εμφανίζονται στα σχήματα 34 και 35. Οι μπλε γραμμές αντιπροσωπεύουν τις τροχιές, ενώ τα κίτρινα σημεία αντιπροσωπεύουν το σημείο εκκίνησης και τα κόκκινα σημεία το τελικό σημείο. Το πρώτο σύμπλεγμα που περιέχει 117 τροχιές απεικονίζεται στο Σχήμα 34. Όπως φαίνεται, οι περισσότερες

τροχιές κατανέμονται σε τουριστικές πόλεις. Άλλωστε, τα περισσότερα από αυτά είναι σύντομα ταξίδια μετ' επιστροφής. Διαπιστώνουν ότι οι τροχιές σε αυτό το σύμπλεγμα δημιουργούνται ως επί το πλείστον από μικρά επιβατηγά σκάφη. Το σύμπλεγμα στο σχήμα 35 περιέχει 180 τροχιές. Άρα, οι (Yao et al., 2018), μπορούν εύκολα να διαπιστώσουν ότι οι περισσότερες από τις τροχιές κατανέμονται στον εσωτερικό ποταμό και οι τροχιές είναι πιο αραιές και μεγαλύτερες από αυτές του πρώτου συμπλέγματος. Όμως, αφού εξέτασαν τις τροχιές των μελών σε αυτό το σύμπλεγμα, διαπίστωσαν ότι οι περισσότερες από αυτές δημιουργήθηκαν από εσωτερικά φορτηγά σκάφη. Τα παραπάνω πειράματα δείχνουν ότι τα συμπλέγματα που δημιουργούνται από την προσέγγισή τους μπορούν να συλλάβουν τα μοτίβα κίνησης των αντικειμένων σε διαφορετικό χρόνο και χώρο. Οι τροχιές σε κάθε ομάδα είναι σημαντικές και μπορούν εύκολα να ερμηνεύσουν κάθε ομάδα αναλύοντας τις τυπικές τροχιές που βρίσκονται σε αυτές. Προηγούμενες μελέτες όπως αυτές των (de Souza et al., 2016) και (Mazzarella et al., 2014) έχουν δείξει ότι διαφορετικοί τύποι σκαφών έχουν διαφορετικά πρότυπα συμπεριφοράς. Ωστόσο, σε αυτήν την εργασία τους, οι συγγραφείς (Yao et al., 2018) προσπαθούν να αναγνωρίσουν τον τύπο του σκάφους χρησιμοποιώντας την ομαδοποίηση της τροχιάς. Λαμβάνουν τα διανύσματα συμπεριφοράς της κίνησης της τροχιάς ενός σκάφους, ως είσοδο του κωδικοποιητή και ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα μεταξύ της εισόδου κωδικοποιητή και της εξόδου αποκωδικοποιητή. Ακολούθως, απέκτησαν το φορέα κινούμενης συμπεριφοράς του σκάφους. Με βάση αυτά τα διανύσματα συμπεριφοράς κινούμενων σκαφών, χρησιμοποίησαν τον αλγόριθμο ομαδοποίησης για να λάβουν τα συμπλέγματα των σκαφών. Ίδανικά, τα σκάφη που ανήκουν σε διαφορετικές ομάδες πρέπει να αποτελούνται από διαφορετικούς τύπους σκαφών. Τα αποτελέσματα ακρίβειας ομαδοποίησης παρουσιάζονται στον Πίνακα 2. Παρόλο που η προσέγγισή τους είναι εντελώς ανεπιτήρητη, εξακολουθούν να έχουν αρκετά καλά ποσοστά ακρίβειας κατά την αναγνώριση σκαφών. Η συνολική ακρίβεια για την αναγνώριση τύπου σκάφους είναι περίπου 78,5%. Ειδικότερα, η ακρίβεια/ανάκληση για το πετρελαιοφόρο και το επιβατηγό σκάφος είναι 0,76/1,0 και 0,87/0,92, αντίστοιχα. Ωστόσο, το αποτέλεσμα του φορτηγού σκάφους είναι μόνο 0,62/0,46. Για να καταλάβουν το λόγο που παρατηρείται αυτό το φαινόμενο, συμβουλευτήκαν τους ειδικούς στον τομέα της ναυτιλίας. Ο λόγος λοιπόν είναι ότι τα φορτηγά σκάφη περιέχουν πολλούς υποτύπους όπως, σκάφος ξηρού φορτίου, σκάφος υγρού φορτίου

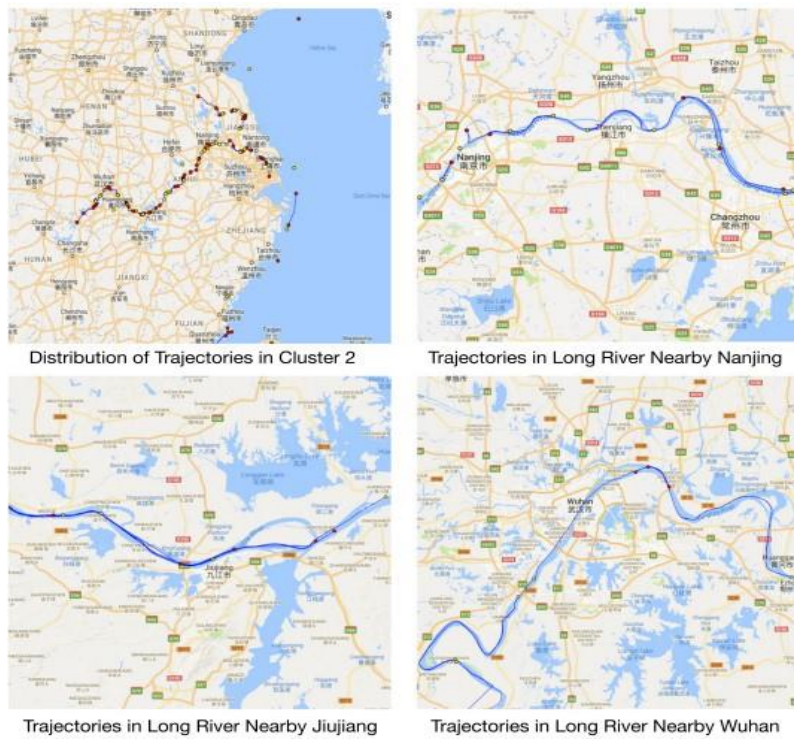
και σκάφος RoRo (Roll-on/roll-off). Αυτοί οι διαφορετικοί τύποι κάνουν μεγάλη διαφορά στα κινούμενα πρότυπα συμπεριφοράς. Ωστόσο, εάν υπάρχουν τέτοιες πληροφορίες υποτύπων στη διαδικασία εκπαίδευσης, η απόδοση του συμπλέγματος αναμένεται να είναι καλύτερη. Σε γενικές γραμμές, αυτό το σύνολο πειραμάτων δείχνει ότι διαφορετικοί τύποι σκαφών έχουν διαφορετικά πρότυπα κινούμενης συμπεριφοράς, επομένως το πλαίσιο των συγγραφέων (Yao et al., 2018) αποτελεί μία εξαιρετική λύση για την καταγραφή τέτοιων μοτίβων.



Σχήμα 33 - Απεικόνιση της μεθόδου ELBOW για να επιλεγεί η K . E_k με την κατάλληλη τιμή K , η οποία σε αυτό το σχήμα πρέπει να είναι το σημείο αγκώνα. Εδώ δίνεται $K = 33$. (Πηγή: Yao et al., 2018)



Σχήμα 34 – Απεικόνιση τροχιών στο Cluster 1 (Πηγή: Yao et al., 2018)



Σχήμα 35 – Απεικόνιση τροχιών στο Cluster 2 (Πηγή: Yao et al., 2018)

	Passenger	Fishing	Cargo	Oil
Total number	50	50	50	50
Precision	46/53=0.87	38/44=0.86	23/37=0.62	50/66=0.76
Recall	46/50=0.92	38/50=0.76	23/50=0.46	50/50=1.0
Overall accuracy:	$(46+38+23+50)/200 = 0.785$			

Πίνακας 2 – Αποτελέσματα ομαδοποίησης τύπου σκαφών (Πηγή: Yao et al., 2018)

Κεφάλαιο 5. Πρακτικό Μέρος – Βελτίωση της εργασίας των Yao et al. (2018) – Συγκριτική Αξιολόγηση στο κομμάτι της ομαδοποίησης

Το πρακτικό μέρος της εργασίας αφορά την μεθοδολογία των Yao et al., (2018), στην οποία επεκτείνουμε ένα μέρος του συνολικού pipeline της προσέγγισης, το οποίο αφορά την διαδικασία της ομαδοποίησης. Αξιολογούμε και συγκρίνουμε τα αποτελέσματα της προσέγγισης αυτής με τη χρήση 4 παραδοσιακών μεθόδων ομαδοποίησης.

Η αρχική εργασία παράγει embeddings από τα δεδομένα των τροχιών και στη συνέχεια πραγματοποιεί συσταδοποίηση. Τα συνθετικά δεδομένα είναι παραγόμενες τροχιές 3 κατηγοριών. Η κάθε προσομοιωμένη τροχιά είναι ένα σύνολο σημείων. Κατά μέσο όρο οι τροχιές έχουν $110,5 \pm 17.4$ σημεία. Οι συντεταγμένες x, y των σημείων κυμαίνονται στο διάστημα $[-20000, 20000]$ μέτρα. Τα δείγματα που προτείνει ο συγγραφέας στο παράδειγμα του στον κώδικα που επισυνάπτει είναι 30 δείγματα. Στην παρούσα εργασία δημιουργήσαμε 300 δείγματα, 100 για κάθε κλάση, για μεγαλύτερη αξιοπιστία των αποτελεσμάτων.

Το τελευταίο βήμα γίνεται με χρήση του αλγορίθμου K-means ($K=3$). Σε αυτή την εργασία θα δοκιμάσουμε εναλλακτικούς αλγορίθμους συσταδοποίησης με σκοπό να βελτιώσουμε την απόδοση του αλγορίθμου.

Δοκιμάσαμε τους αλγόριθμους DBSCAN, Spectral, Ιεραρχικός συσσωρευτικός με τα κριτήρια single, complete, ward. Πρόκειται για έναν αλγόριθμο που βασίζεται στην πυκνότητα (DBSCAN), ενός διαχωριστικού (spectral) και τριών ιεραρχικών αλγορίθμων. Οι αλγόριθμοι υλοποιήθηκαν και εκτελέστηκαν με τις default παραμέτρους που ορίζονται από τους συγγραφείς. Επειδή ο κάθε αλγόριθμος έχει διαφορετική λογική, θα προκύπτουν διαφορετικές ομαδοποιήσεις. Οπότε ο σκοπός μας είναι να διαπιστώσουμε αν κάποιος αλγόριθμος μπορεί να εκμεταλλευθεί καλύτερα τα χαρακτηριστικά αυτών των δεδομένων.

Αξιολογήσαμε την ομαδοποίηση του κάθε αλγορίθμου και τα αποτελέσματα είναι τα παρακάτω. Στον Πίνακα 3 φαίνεται η ακρίβεια κάθε αλγορίθμου σε κάθε μια από τις τρεις κατηγορίες (δηλαδή το ποσοστό των τροχιών που είναι στην ίδια ομάδα) και συνολικά (ο μέσος όρος των τριών κατηγοριών). Παρατηρούμε ότι ο Ιεραρχικός αλγόριθμος με κριτήριο single linkage έχει την καλύτερη απόδοση.

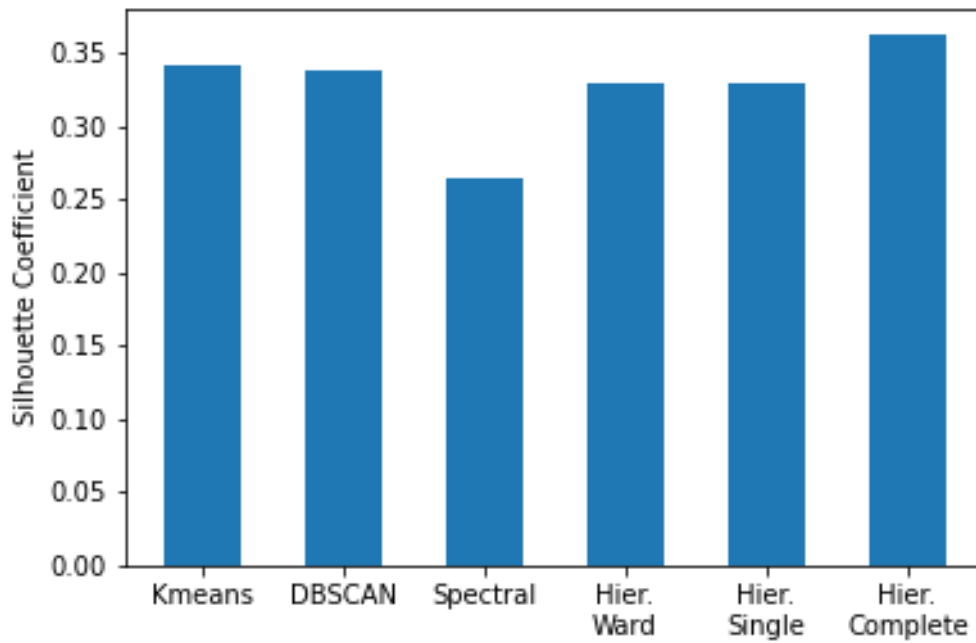
Πίνακας 3 - Ακρίβεια αλγορίθμων ανά κατηγορία τροχιών και συνολικά σε ένα πίνακα

Αλγόριθμος	Ευθείες (%)	Κυκλικές (%)	Καμπύλες (%)	Συνολικά
K-means	90	100	70	86.6
DBSCAN	100	50	100	83.3
Spectral	100	50	100	83.3
Hier. Ward	80	100	70	83.3
Hier. Single	100	80	100	93.3
Hier. Complete	100	70	100	90

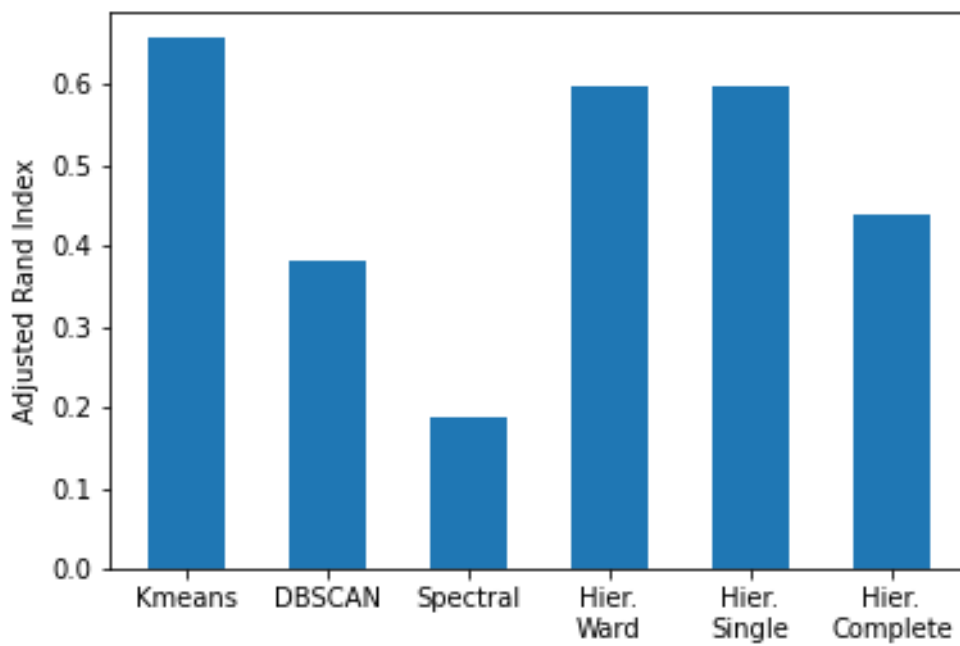
Για την αξιολόγηση χρησιμοποιήσαμε το εσωτερικό μέτρο Silhouette (Σχήμα 31) και τα εξωτερικά μέτρα Adjust Rand Index (ARI) (Σχήμα 32), Adjusted Mutual Information (AMI) (Σχήμα 33), Accuracy (Σχήμα 34) και V-measure (Σχήμα 35). Επιλέξαμε τόσο τα εσωτερικά μέτρα τα οποία αξιολογούν την ομαδοποίηση με κριτήριο τις αποστάσεις των σημείων μεταξύ τους όσο και τα εξωτερικά που λαμβάνουν υπόψη τους την κλάση του κάθε σημείου και κατά πόσο η κάθε ομάδα αντιπροσωπεύει κάποια από τις κλάσεις αυτές.

Με την εσωτερική μετρική Silhouette, καλύτερος αναδεικνύεται ο Ιεραρχικός με κριτήριο complete, αλλά οι υπόλοιποι έχουν πολύ μικρή διαφορά στην απόδοση, εκτός του Spectral, που έχει αρκετά μικρότερη. Με τις εξωτερικές μετρικές ARI, AMI και V-measure τα αποτελέσματα είναι πολύ παρόμοια, καθώς αναδεικνύεται ως καλύτερος ο K-means, μετά οι ιεραρχικοί και στη συνέχεια ακολουθούν ο DBSCAN και ο Spectral. Με τη μετρική Accuracy, όπως είδαμε και αναλυτικά στον Πίνακα 3, ο καλύτερος αλγόριθμος είναι ο Ιεραρχικός με κριτήριο single, μετά ο Ιεραρχικός με complete, τρίτος ο K-means και μετά οι άλλοι αλγόριθμοι με μικρή διαφορά.

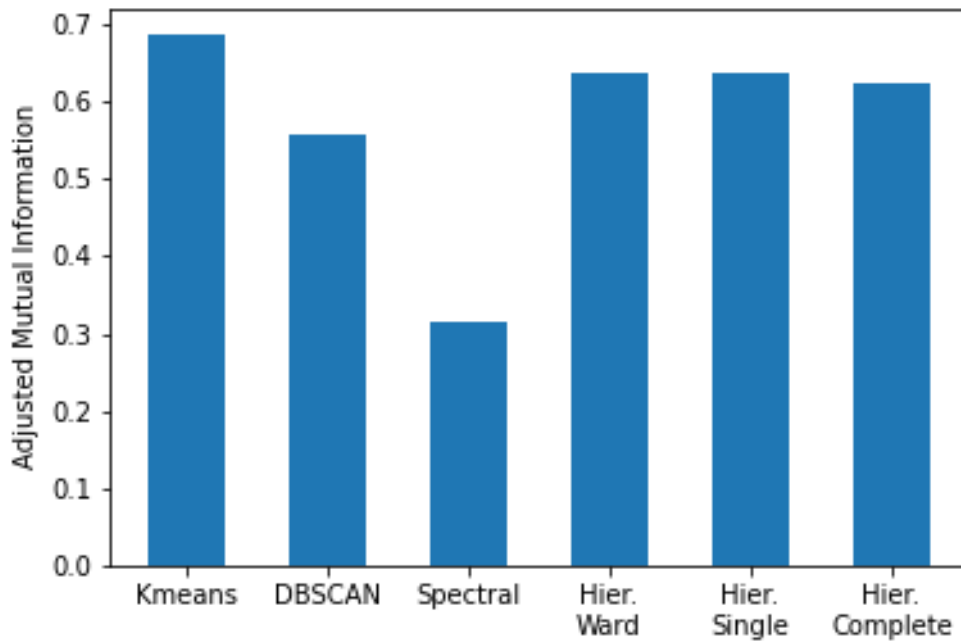
Οπότε διαπιστώνουμε ότι ανάλογα με τη μετρική, αλλάζει ποιος είναι ο καλύτερος αλγόριθμος. Με το εσωτερικό κριτήριο, το οποίο δείχνει πόσο καλά διαχωρίζονται οι ομάδες χωρίς να ξέρουμε την πραγματική κατηγορία τους, προηγείται ο Ιεραρχικός με κριτήριο complete. Ο K-means αποδεικνύεται ότι είναι η καλύτερη επιλογή με τα εξωτερικά μέτρα ARI, AMI, V-measure, αλλά με βάση τη μετρική Accuracy, στην οποία δίνεται περισσότερη έμφαση στην εργασία του Yao, ο Ιεραρχικός με κριτήριο single αναδεικνύεται ο καλύτερος.



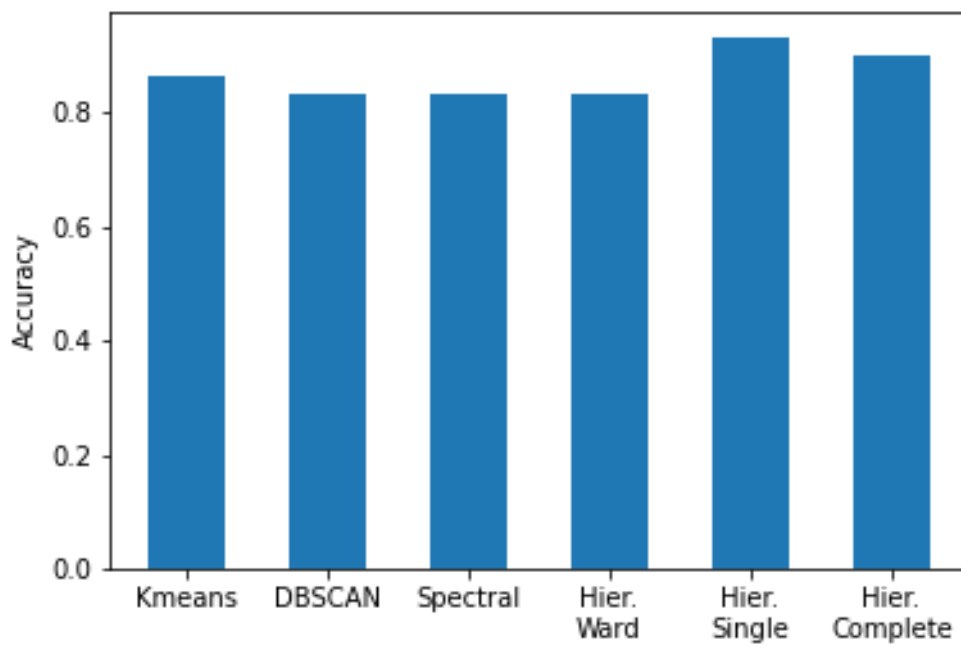
Σχήμα 36 - Αξιολόγηση ομαδοποίησης με την μετρική Silhouette



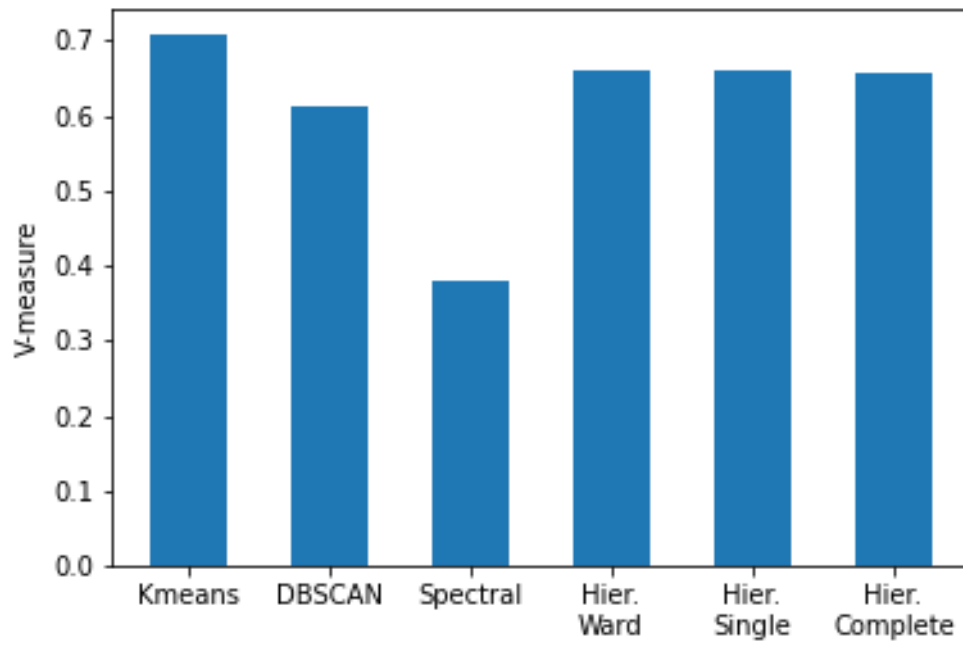
Σχήμα 37 - Αξιολόγηση ομαδοποίησης με την μετρική ARI



Σχήμα 38 - Αξιολόγηση ομαδοποίησης με την μετρική AMI



Σχήμα 39 - Αξιολόγηση ομαδοποίησης με την μετρική Accuracy



Σχήμα 40 - Αξιολόγηση ομαδοποίησης με την μετρική V-measure

Συμπεράσματα

Η ανάλυση των δεδομένων κίνησης αποτελεί μια ιδιαίτερα σύνθετη διαδικασία τόσο λόγω των υπεισερχομένων παραμέτρων όσο και τις δυσκολίας υπολογισμού που στην τρέχουσα περίοδο έχει να αντιμετωπίσει και την διαχείριση του υπερβολικά μεγάλου όγκου δεδομένων. Παρόλα αυτά θεωρείται, και πιθανότατα είναι η αιχμή της τεχνολογίας καθώς κάθε εφαρμογή κινητού τηλεφώνου που ασχολείται με την ανθρώπινη δραστηριότητα κάνει καταγραφή της θέσης του. Πέραν όμως αυτών των εφαρμογών, υπάρχουν εφαρμογές και στην ζωολογία που μελετούν την μετανάστευση ή την απλή κίνηση ζωντανών οργανισμών.

Ενώ η ανάλυση των δεδομένων κίνησης με βάση συγκεκριμένα μέτρα στατιστικής, όπως π.χ. μέση ταχύτητα κίνησης, απόκλιση από την τροχιά κ.α. αποτελούν μια σχετικά απλή διαδικασία, η συσταδοποίηση των δεδομένων παρουσιάζει μια ιδιαίτερη δυσκολία. Η συσταδοποίηση δεδομένων κίνησης διαφέρει σημαντικά από τις κλασικές μεθόδους συσταδοποίησης καθώς οι υπολογισμοί γίνονται με βάση την απόσταση μεταξύ των σημείων και όχι ανά παρατήρηση. Αυτό σημαίνει ότι ο τρόπος υπολογισμού της απόστασης μπορεί να έχει και διαφορετικά αποτελέσματα στον ίδιο αλγόριθμο συσταδοποίησης.

Στην παρούσα εργασία υλοποιήσαμε την εργασία των Yao et al., (2018) η οποία παράγει embeddings από τα δεδομένα των τροχιών με χρήση αλγορίθμων βαθιάς μάθησης (DNN) και στη συνέχεια πραγματοποιεί συσταδοποίηση με τον k-means. Μετέπειτα προσπαθήσαμε να βελτιώσουμε τα αποτελέσματα της πρωτότυπης εργασίας μέσα από την εφαρμογή άλλων αλγορίθμων. Η διαίσθηση μας ήταν ότι με τον k-means που είναι ένας απλός διαχωριστικός αλγόριθμος, η μέθοδος δεν είχε βέλτιστα αποτελέσματα. Κάτι που επιβεβαιώθηκε, μιας και ο ιεραρχικός συσσωρευτικός αλγόριθμος με κριτήριο το single linkage είχε το καλύτερο αποτέλεσμα με την μετρική «Ακρίβεια» για την ταξινόμηση των διαφορετικών κατηγοριών τροχιών.

Παράρτημα - Εγκατάσταση

Η εγκατάσταση και εκτέλεση του κώδικα έγινε σε Ubuntu 18.04, το οποίο εκτελούνταν σε εικονική μηχανή Oracle Virtual Box. Σημειώνεται ότι αρχικά δοκιμάσαμε να εργαστούμε σε Windows, ωστόσο λόγω διάφορων προβλημάτων κατά την μεταγλώττιση (compile), στραφήκαμε σε Linux. Οπότε, εγκαταστήσαμε το Anaconda (anaconda.com), το οποίο προσφέρει διάφορα εργαλεία για τη χρήση Python. Τα απαιτούμενα πακέτα επειδή έχουν δημιουργηθεί πριν κάποια χρόνια έχουν γραφτεί σε Python 2. Οπότε δημιουργούμε ένα εικονικό περιβάλλον με Python 2.7 με την εντολή:

- `conda create -n py27 python=2.7`

Ύστερα, ενεργοποιούμε το εικονικό περιβάλλον:

- `conda activate py27`

Μετά, εγκαθιστούμε τα προαπαιτούμενα πακέτα του πακέτου traj-dist (https://github.com/maikol-solis/trajectory_distance):

- `pip install numpy cython scipy geohash pandas shapely`

Στη συνέχεια μετακινούμαστε στο φάκελο των πηγαίων αρχείων του πακέτου και τα μεταγλωττίζουμε:

- `python setup.py build_ext -inplace`

Όταν τελειώσει η προηγούμενη διαδικασία και δημιουργηθούν τα απαραίτητα αρχεία, το εγκαθιστούμε:

- `pip install .`

Για να ελέγξουμε ότι όλα έγιναν σωστά, σε ένα τερματικό Python δίνουμε την παρακάτω εντολή, η οποία πρέπει να ολοκληρωθεί χωρίς σφάλματα:

- `import traj_dist.distance as tdist`

Στη συνέχεια, εγκαθιστούμε τα προαπαιτούμενα πακέτα για να εκτελεστεί ο αλγόριθμος trajectory2vec (<https://github.com/yaodi833/trajectory2vec>):

- `pip install sklearn`
- `pip install tensorflow==0.12.0rc0`

Σε αυτό το σημείο όλα είναι έτοιμα για να εκτελεστούν τα αρχεία του αλγορίθμου

- `trajectory2vec.`

Βιβλιογραφικές αναφορές

- Afshine, A., Shervine, A., (2020). Recurrent Neural Networks cheatsheet. Ανακτήθηκε 1/3/2020 <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks>
- Aggarwal, C.C. (2018). *Neural Networks and Deep Learning*. Switzerland: Springer.
- Anjum, N., & Cavallaro, A. (2008). Multifeature object trajectory clustering for video analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(11), 1555-1564.
- Ankerst, M., Breunig, M., Kriegel, H.-P. & Sander, J. (1999). Optics: Ordering points to identify the clustering structure. In *Proceedings of ACM SIGMOD international conference on management of data (SIGMOD'99)* Philadelphia, Pennsylvania. New York: ACM.
- Balouji, E., Gu, I.Y.H., Bollen, M.H.J., Bagheri and M. Nazari, A., (2018). A LSTM-based deep learning method with application to voltage dip classification. 18th International Conference on Harmonics and Quality of Power (ICHQP), Ljubljana, Slovenia, pp. 1-5, doi: 10.1109/ICHQP.2018.8378893.
- Besse, P. C., Guillouet, B., Loubes, J. M., & Royer, F. (2016). Review and perspective for distance-based clustering of vehicle trajectories. *IEEE Transactions on Intelligent Transportation Systems*, 17(11), 3306-3317.
- Bian, J., Tian, D., Tang, Y., & Tao, D. (2018). A survey on trajectory clustering analysis. *arXiv preprint arXiv:1802.06971*.
- Boyle, J., Nawaz, T., & Ferryman, J. (2017, August). Deep trajectory representation-based clustering for motion pattern extraction in videos. In *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1-6). IEEE.
- Cao, H., Mamoulis, N. & Cheung, D.W. (2005). Mining frequent spatiotemporal sequential patterns. In: *Proceedings of ICDM*.
- Cao, Y., Yuan, J., Xiao, S., Xie, Q. (2019). *TPM: A GPS-based Trajectory Pattern Mining System*.

- Chang, C. & Zhou, B. (2009). Multi-granularity Visualization of Trajectory Clusters Using Sub-trajectory Clustering, *IEEE International Conference on Data Mining Workshops, Miami, FL*, pp. 577-582, doi: 10.1109/ICDMW.2009.24.
- Chung, Y. A., Wu, C. C., Shen, C. H., Lee, H. Y., & Lee, L. S. (2016). Unsupervised learning of audio segment representations using sequence-to-sequence recurrent neural networks. In *Proc. Interspeech* (pp. 765-769).
- Dai, A. M., & Le, Q. V. (2015). Semi-supervised sequence learning. *Advances in neural information processing systems*, 28, 3079-3087.
- de Souza, E. N., Boerder, K., Matwin, S., & Worm, B. (2016). Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PloS one*, 11(7), e0158248.
- Feng, Z., & Zhu, Y. (2016). A survey on trajectory data mining: Techniques and applications. *IEEE Access*, 4, 2056-2067.
- Gaffney, S. & Padhraic, S. (1999). *Trajectory Clustering with Mixtures of Regression Models*. doi: 10.1145/312129.312198.
- GG, (2020). Understanding of LSTM Networks. Ανακτήθηκε 18/2/2021 από <https://www.geeksforgeeks.org/understanding-of-lstm-networks/>
- Giannotti, F., Mazzoni, A., Puntoni, S. & Renso, C. (2005). Synthetic generation of cellular network positioning data. *Technical report*, ISTI-CNR.
- Giannotti, F., Pappalardo, L., Pedreschi, D. & Wang, D. (2013). A Complexity Science Perspective on Human Mobility, in *Mobility Data*, C. Renso, S. Spaccapietra, and E. Zimanyi, Eds. Cambridge: Cambridge University Press, pp. 297–314.
- Gruner, R., Henzler, P., Hinz, G., Eckstein, C., & Knoll, A. (2017, June). Spatiotemporal representation of driving scenarios and classification using neural networks. In *2017 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1782-1788). IEEE.
- Han, B., Liu, L. & Omiecinski, E., (2012). NEAT: Road Network Aware Trajectory Clustering. *IEEE 32nd International Conference on Distributed Computing Systems*, Macau, 2012, pp. 142-151, doi: 10.1109/ICDCS.2012.31.

- Hardesty, L. (2017). Explained: Neural networks. Ανακτήθηκε από <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>
- Harmening, N., Biloš, M., & Günnemann, S. (2020). Deep representation learning and clustering of traffic scenarios. *arXiv preprint arXiv:2007.07740*
- Hung C.-C., Peng W.-C. & Lee W.-C., (2015). Clustering and aggregating clues of trajectories for mining trajectory patterns and routes. *The VLDB Journal*, 24:169–192
- IBM, (2020). Recurrent Neural Networks. Ανακτήθηκε 1/3/2020 από <https://www.ibm.com/cloud/learn/recurrent-neural-networks>
- Ihaka, R. & Gentleman, R. (1996). A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314.
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732).
- Kelly, J., Saripalli, S., Sukhatme, G., (2007). Combined Visual and Inertial Navigation for an Unmanned Aerial Vehicle. 10.1007/978-3-540-75404-6_24.
- Kong, X., Xia, F., Ning, Z., Rahim, A., Cai, Y., Gao, Z. & Ma, J. (2018). Mobility dataset generation for vehicular social networks based on floating car data, *IEEE Transactions on Vehicular Technology*, 67(5), 3874–3886.
- Lea, C., Vidal, R., Reiter, A., & Hager, G. D. (2016, October). Temporal convolutional networks: A unified approach to action segmentation. In *European Conference on Computer Vision* (pp. 47-54). Springer, Cham.
- Lee W.-C. & Krumm, J. (2011). *Trajectory Preprocessing In Computing with Spatial Trajectories* (Edts Yu Zheng, Y., Zhou, X.). USA, Springer.
- Leffondré K, Abrahamowicz M, Regeasse A, Hawker GA, Badley EM, McCusker J, Belzile E. (2004). Statistical measures were proposed for identifying longitudinal patterns of change in quantitative health indicators. *J Clin Epidemiol.*, 57(10):1049-62. doi: 10.1016/j.jclinepi.2004.02.012. PMID: 15528056.

- Lyon, M. & Hafner, K., (1998). *Where Wizards Stay Up Late: The Origins Of The Internet*. USA: Simon & Schuster, Inc.
- Ma, X., Liu, C., Wen, H., Wang, Y. & Wu, Y.J., (2017). Understanding commuting patterns using transit smart card data, *Journal of Transport Geography*, 58, 135–145,
- Maratnia, N. & de By, R. (2004). Spatio-Temporal Compression Techniques for Moving Point Objects. In: *International Conference on Extending Database Technology (EDBT)*, 765–782
- Mazimpaka, J.D. & Timpf, S., (2016). Trajectory data mining: A review of methods and applications, *Journal of Spatial Information Science*, no. 13.
- Mazzarella, F., Vespe, M., Damalas, D., & Osio, G. (2014, July). Discovering vessel activities at sea using AIS data: Mapping of fishing footprints. In *17th International conference on information fusion (FUSION)* (pp. 1-7). IEEE.
- McCullough, S.W. & Pitts, H.W. (1944). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5 pp.115-133
- Min, E., Guo, X., Liu, Q., Zhang, G., Cui, J., & Long, J. (2018). A survey of clustering with deep learning: From the perspective of network architecture. *IEEE Access*, 6, 39501-39514.
- Minsky, M. & Papert, S.A. (2017). *Perceptrons: An Introduction to Computational Geometry*. USA: The MIT Press.
- Pelletier, M.P., TrÃpanier, M., Morency, C., (2011). “Smart card data use in public transit: A literature review,” *Transportation Research Part C*, 19, (4), 557–568.
- Nanni, M. & Pedreschi, D. (2006). Time-focused clustering of trajectories of moving objects. *J Intell Inf Syst* 27, 267–289. Ανακτήθηκε από <https://doi.org/10.1007/s10844-006-9953-7>
- Olive, X., Basora, L., Viry, B., & Alligier, R. (2020, September). Deep trajectory clustering with autoencoders. In *Proceedings of the International Conference on Research in Air Transportation*.

- Pelekis, N., Kopanakis, I., Kotsifakos, E. E., Frentzos, E., & Theodoridis, Y. (2011). Clustering uncertain trajectories. *Knowledge and information systems*, 28(1), 117-147.
- Qiao, Y. Cheng, Y., Yang, J. Liu, J. & Kato, N. (2017). A Mobility Analytical Framework for Big Mobile Data in Densely Populated Area, *IEEE Transactions on Vehicular Technology*, 66(2), 1443–1455.
- Renso, C., Baglioni, M., de Macedo, J.A.F., Trasarti, R. & Wachowicz, M. (2013). How you move reveals who you are: Understanding human behavior by analyzing trajectory data, *Knowledge and Information Systems*, 37(2), 331–362.
- Rosenblat, F. (1957). *The perceptron. A perceiving and recognizing automation*. Ανακτήθηκε από <https://blogs.umass.edu/brain-wars/files/2016/03/rosenblatt-1957.pdf>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- Shaw, S.L., Tsou, M.H. & Ye, X. (2016). Editorial: Human dynamics in the mobile and big data era. *International Journal of Geographical Information Science*, 30(6), 1687–1693.
- Sylvestre MP & Vatnik, D. (2014). Using traj Package to Identify Clusters of Longitudinal Trajectories. Ανακτήθηκε από <https://cran.r-project.org/web/packages/traj/vignettes/trajVignette.pdf>
- Tampakis, P., Pelekis, N., Doulkeridis, C., Theodoridis, Y.,. (2019). Scalable Distributed Subtrajectory Clustering.
- Wang, D., Miwa, T., Morikawa, T.,. (2020). Big Trajectory Data Mining: A Survey of Methods, Applications, and Services. *Sensors*, 20(16):4571. <https://doi.org/10.3390/s20164571>
- Wang, S., Bao, Z., Culpepper, J., Sellis, T., Qin, X., (2019). Fast large-scale trajectory clustering. *Proceedings of the VLDB Endowment*. 13. 29-42. 10.14778/3357377.3357380.

- Xavier, O., Luis B., Benoit, V., Richard, A., (2020). Deep Trajectory Clustering with Autoencoders. ICRAAT 2020, 9th International Conference for Research in Air Transportation, Sep 2020, Tampa, United States. ffhal-02916241f
- Xu, H., Zhou, Y., Lin, W. & Zha, H. (2015). Unsupervised Trajectory Clustering via Adaptive Multi-kernel-Based Shrinkage, *IEEE International Conference on Computer Vision (ICCV)*, Santiago, pp. 4328-4336, doi: 10.1109/ICCV.2015.492.
- Yao, D., Zhang, C., Zhu, Z., Hu, Q., Wang, Z., Huang, J., & Bi, J. (2018). Learning deep representation for trajectory clustering. *Expert Systems*, 35(2), e12252.
- Yue, M., Li, Y., Yang, H., Ahuja, R., Chiang, Y-Y., Shahabi, C., (2019). DETECT: Deep Trajectory Clustering for Mobility-Behavior Analysis. 988-997. 10.1109/BigData47090.2019.9006561.
- Zhang, Y., Hare, J., & Prugel-Bennett, A. (2019). Deep set prediction networks. *Advances in Neural Information Processing Systems*, 32, 3212-3222.
- Zheng, Y., (2015). Trajectory Data Mining: An Overview. *ACM Transactions on Intelligent Systems and Technology*, 6(3), 1–41.
- Zhou, J., Murphy, E. & Long, Y. (2014). Commuting efficiency in the Beijing metropolitan area: An exploration combining smartcard and travel survey data. *Journal of Transport Geography*, 41, 175–183.