



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής

Πρόγραμμα Μεταπτυχιακών Σπουδών

«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Διαχείριση σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων Management of moving objects' semantic trajectories
Όνοματεπώνυμο Φοιτητή	Θεόδωρος Γουμενίδης
Πατρώνυμο	Χρήστος
Αριθμός Μητρώου	ΜΠΣΠ/ 11040
Επιβλέπων	Νίκος Πελέκης, Λέκτορας

Ημερομηνία Παράδοσης **Ιούλιος 2013**

Πανεπιστήμιο Πειραιώς

Πανεπιστήμιο Πειραιώς

Τριμελής Εξεταστική Επιτροπή

Νίκος Πελέκης
Λέκτορας

Ιωάννης Θεοδωρίδης
Καθηγητής

Ιωάννης Σίσκος
Καθηγητής

Πανεπιστήμιο Πειραιώς

Περίληψη

Καθώς ολοένα και περισσότερες συσκευές και υπηρεσίες παρέχουν τη δυνατότητα εντοπισμού θέσης, η παραγωγή δεδομένων κίνησης αυξάνεται με γεωμετρικούς ρυθμούς. Οι απαιτήσεις που προκύπτουν από την τάση αυτή οδήγησαν στην ανάπτυξη των βάσεων δεδομένων κινούμενων αντικειμένων (MOD) και κατ' επέκταση στις αποθήκες δεδομένων τροχιών (TDW). Ωστόσο, οι προκλήσεις για την αποτελεσματικότερη διαχείριση και εκμετάλλευση τους παραμένουν, και γι αυτό μια νέα αντιμετώπιση που αφορά στην ενσωμάτωση της σημασιολογικής στην χωροχρονική πληροφορία κερδίζει χώρο, δημιουργώντας μια νέα κατηγορία δεδομένων, τις σημασιολογικά επαυξημένες τροχιές. Η παρούσα λοιπόν, εργασία βασίζεται στο σύστημα HERMES MOD και σε περαιτέρω λειτουργίες του προς αυτή την κατεύθυνση, εξετάζει τις δυνατότητες που παρέχονται ως προς την διαχείριση και την εκμετάλλευση τους. Για τον σκοπό αυτό εφαρμόζεται ένα πλήθος σεναρίων με την μορφή ερωτημάτων πάνω σε δοθέν σύνολο πραγματικών δεδομένων από το οποίο με την κατάλληλη επεξεργασία προκύπτουν σημασιολογικά επαυξημένες τροχιές. Επίσης, ως προς την κατεύθυνση αυτή αναπτύχθηκαν διάφορες επεκτάσεις, οι οποίες αφορούν τόσο τον χειρισμό όσο και την παραγωγή γνώσης από τέτοιου είδους δεδομένα. Παρουσίαση του ψευδο-κώδικα και εξέταση της λειτουργικότητας τους γίνεται σε κεφάλαια που ακολουθούν.

Λέξεις Κλειδιά: σημασιολογικά επαυξημένες τροχιές, βάσεις δεδομένων τροχιών, HERMES

Abstract

As more and more devices and services enable positioning capability, production data traffic is growing exponentially fast. The requirements arising from this trend led to the development of Moving Objects Databases (MOD) and by extension in the Trajectory Data Warehouses (TDW). However, the challenges for more effective management and exploitation remain, and that is why a new approach regarding the integration of semantic on spatiotemporal information is gaining space, creating a new category of data, the semantically augmented orbits. Therefore, this project, based on HERMES MOD system and further operational functions in this direction, examines the features provided as far as their management and exploitation is concerned. For this purpose, a plurality of scenarios is applied in the form of queries on a given set of actual data from which augmented semantic paths result after the appropriate processing. Furthermore, various extensions were developed in this direction, which affect both the treatment and the production of knowledge from such data. Presentation of pseudo-code and testing of their functionality is given in the following chapters.

Keywords: semantic trajectories, semantic trajectory databases, HERMES

Πίνακας περιεχομένων

1	Εισαγωγή	8
2	Θεωρητικό Υπόβαθρο	9
2.1	Τροχιές κινουμένων αντικειμένων (trajectories)	9
2.2	Σημασιολογικά επαυξημένες τροχιές (semantic trajectories)	10
2.2.1	Περιγραφή	10
2.2.2	Μοντέλα αναπαράστασης-επεξεργασίας σημασιολογικά επαυξημένων τροχιών	11
2.2.3	Σημασιολογικός εμπλουτισμός της τροχιάς	12
2.3	Βάσεις κινούμενων αντικειμένων (MOD) - HERMES	13
3	Ανάλυση Απαιτήσεων	15
3.1	Κριτήρια αξιολόγησης	15
3.2	Προσδοκίες χρήσης	16
4	Σχεδιασμός	17
4.1	Υπάρχουσα δομή	17
4.1.1	Σημασιολογικά επαυξημένες τροχιές στο HERMES MOD	17
4.1.2	Αποθήκη δεδομένων σημασιολογικά επαυξημένων τροχιών στο HERMES MOD	18
4.2	Σχεδιασμός ερωτημάτων	19
4.3	Επεκτάσεις	19
4.3.1	Απεικόνιση σημασιολογικά επαυξημένης τροχιάς	19
4.3.2	Εύρεση K πλησιέστερων σημείων ενδιαφέροντος ενός επεισοδίου	19

4.3.3	Εύρεση πιθανοτήτων για τα Κ πλησιέστερα σημεία ενδιαφέροντος ενός επεισοδίου.	20
4.3.4	Δημιουργία μήτρας προέλευσης-προορισμού.	20
5	Παράδειγμα χρήσης.	23
5.1	Το δοθέν σύνολο δεδομένων	23
5.2	Εξερεύνηση δυνατοτήτων HERMES MOD	26
5.3	Ερωτήματα στο HERMES MOD	27
5.3.1	Εφαρμογή τελεστών για τον υπολογισμό του αριθμού των επεισοδίων	27
5.3.2	Χρήση τελεστή για την εύρεση επεισοδίων	28
5.3.3	Σύνθετη χρήση τελεστών πάνω στην σημασιολογική πληροφορία	29
5.3.4	Περιορισμός στη χωρική διάσταση, φιλτράρισμα με βάση τη σημασιολογική πληροφορία κειμένου και οπτικοποίηση των αποτελεσμάτων	29
5.3.5	<i>Ερώτημα χρονικού εύρους και φίλτρου στην σημασιολογική πληροφορία με χρήση ευρετηρίου.</i>	33
5.3.6	<i>Ερώτημα χωρο-χρονικού εύρους και φίλτρου στη σημασιολογική πληροφορία με χρήση ευρετηρίου.</i>	34
5.3.7	<i>Ερώτημα χωρο - χρονικού εύρους διασταύρωσης και φίλτρου στη σημασιολογική πληροφορία με χρήση ευρετηρίου καθώς και εφαρμογή χρονικού περιορισμού.</i>	35
5.3.8	Εφαρμογή τελεστών για την εύρεση επεισοδίων τα οποία επικαλύπτονται με χρονική περίοδο.	38

5.3.9	Ερωτήματα εύρους σε ευρετήριο για την εύρεση μοτίβων της μορφής “από-προς”	39
5.3.10	Η αποθήκη δεδομένων σημασιολογικά επαυξημένων τροχιών 42	
5.3.11	Εφαρμογή ερωτημάτων στην αποθήκη δεδομένων σημασιολογικά επαυξημένων τροχιών	42
5.4	Εφαρμογή Επεκτάσεων	44
5.4.1	Απεικόνιση σημασιολογικά επαυξημένης τροχιάς	44
5.4.2	Εύρεση K πλησιέστερων σημείων ενδιαφέροντος	45
5.4.3	Εύρεση πιθανοτήτων για τα K πλησιέστερα σημεία ενδιαφέροντος ενός επεισοδίου	46
5.4.4	Δημιουργία μήτρας προέλευσης-προορισμού	47
6	Συμπεράσματα	50
7	Βιβλιογραφία	51

Πίνακας εικόνων

Εικόνα 1: (α) Απεικόνιση της τροχιάς σαν συνεχές φαινόμενο και (β) απεικόνιση της σαν μια πεπερασμένη ακολουθία χρονοσημασμένων σημείων	9
Εικόνα 2: Απεικόνιση μιας τροχιάς με την μορφή ακολουθίας επεισοδίων.....	11
Εικόνα 3: Απεικόνιση σημασιολογικά επαυξημένης τροχιάς.....	11
Εικόνα 4: (1) μια απλή τροχιά και (2) μια απλή σημασιολογικά επαυξημένη τροχιά.....	16
Εικόνα 5: Σχηματική απεικόνιση μέρους των αντικειμένων της επέκτασης του HERMES.....	18
Εικόνα 6: Σχηματική απεικόνιση της αποθήκης δεδομένων για σημασιολογικά επαυξημένες τροχιές στο HERMES.....	19
Εικόνα 7: Απεικόνιση μιας μήτρας προέλευσης-προορισμού	21
Εικόνα 8: Αλγόριθμος <code>construct_odmatrix</code>	21
Εικόνα 9: Αλγόριθμος <code>construct_dw_odmatrix</code>	22
Εικόνα 10: Γραφήματα δεδομένων των ανακατασκευασμένων τροχιών από <code>gps</code> εγγραφές.....	24
Εικόνα 11: Κατανομή των STOP επεισοδίων ως προς τον σκοπό στον προορισμό.....	25
Εικόνα 12: Κατανομή των MOVE επεισοδίων ως προς το χρησιμοποιηθέν μεταφορικό μέσο.....	25
Εικόνα 13: Αποτελέσματα ερωτήματος υπολογισμού αριθμού επεισοδίων	28
Εικόνα 14: Αποτελέσματα ερωτήματος.....	28
Εικόνα 15: Αποτελέσματα ερωτήματος με χρήση φίλτρου πολλαπλών ετικετών.....	29
Εικόνα 16: Οπτικοποίηση αποτελεσμάτων. Στις εικόνες φαίνονται, με μπλε χρώμα, το <code>semantic trajectory</code> στο σύνολο του και με κόκκινο χρώμα τα επεισόδια που επιστρέφει το ερώτημα.....	31
Εικόνα 17: Οπτικοποίηση αποτελεσμάτων. Στις εικόνες φαίνονται, με μπλε χρώμα, το <code>semantic trajectory</code> στο σύνολο του. Με κόκκινο χρώμα, τα επεισόδια που επιστρέφει το ερώτημα.....	33
Εικόνα 18: Αποτελέσματα ερωτήματος χρονικού εύρους	34
Εικόνα 19: Αποτελέσματα ερωτήματος χωροχρονικού εύρους.....	35

Εικόνα 20: Στην εικόνα φαίνονται, με μπλε χρώμα, τα επεισόδια με ετικέτα ‘STOP’ της δοθείσας χρονικής περιόδου και γεωμετρικού περιορισμού (περίγραμμα με πράσινο χρώμα). Με κόκκινο χρώμα, το τμήμα του επεισοδίου που βρίσκεται στην χρονική περίοδο που δηλώνετε κατά την χρήση του τελεστή at_period.....	38
Εικόνα 21: Αποτελέσματα ερωτήματος.....	38
Εικόνα 22: Αποτελέσματα ερωτήματος.....	39
Εικόνα 23: Αποτελέσματα ερωτήματος.....	40
Εικόνα 24: Στην εικόνα φαίνονται - με μπλε χρώμα - όλα τα επεισόδια με ετικέτα ‘MOVE’, τα οποία ξεκινούν από επεισόδιο με ετικέτα ‘STOP’ με δοθέν χωροχρονικό περιορισμό και καταλήγουν σε επεισόδιο με ετικέτα ‘STOP’. Ο δοθέν χωρικός περιορισμός φαίνεται με πράσινο περίγραμμα.	41
Εικόνα 25: Εύρεση συνολικής διάρκειας και απόστασης ανά χρονική περίοδο και φύλο όσων κινήθηκαν με τρένο.	43
Εικόνα 26: Αριθμός κινούμενων αντικειμένων προς/σε προορισμό την ίδια χρονική περίοδο.....	44
Εικόνα 27: Χρήση προσθήκης για την απεικόνιση σημασιολογικής τροχιάς. Φαίνονται τα MBRs των επεισοδίων με άσπρο περίγραμμα, η πληροφορία καθώς και οι υπο-τροχιές με μπλε χρώμα.	45
Εικόνα 28: Αποτελέσματα εύρεσης $K(=10)$ πλησιέστερων σημείων ενδιαφέροντος (POI), τα οποία βρίσκονται εντός του MBR (δείχνετε με άσπρο περίγραμμα) ενός επεισοδίου.	46
Εικόνα 29: Αποτελέσματα εύρεσης των πιθανοτήτων των $K(=10)$ πλησιέστερων σημείων ενδιαφέροντος (POI), τα οποία βρίσκονται εντός του MBR (δείχνετε με άσπρο περίγραμμα) ενός επεισοδίου.	47
Εικόνα 30: Η παραγόμενη μήτρα προέλευσης-προορισμού από σημασιολογικά επαυξημένες τροχιές.	48

1 Εισαγωγή

Στη σημερινή εποχή υπάρχει μια αυξανόμενη τάση ως προς τη χρήση συσκευών και τεχνολογιών που παρέχουν τη δυνατότητα εντοπισμού και καταγραφής της θέσης. Τέτοιες συσκευές μπορεί να είναι ένα κινητό τηλέφωνο, ένας GPS δέκτης ή ένα δίκτυο αισθητήρων κ.α. Το γεγονός αυτό έχει ως αποτέλεσμα τη δημιουργία μιας τεράστιας ποσότητας χρονοσημασμένων στιγμάτων, τα οποία αποτελούν τα *δεδομένα κίνησης*. Οι διαδικασίες παραγωγής, μετάδοσης, αποθήκευσης, απεικόνισης, ανασύστασης και επεξεργασίας είναι προβληματικές κι έχουν γίνει κατά καιρούς ποικίλες σχετικές προτάσεις.

Εκτός όμως από τις προκλήσεις ως προς τον χειρισμό των δεδομένων αυτών σημαντική πρόκληση αποτελεί και η αξιοποίηση τους με τη χρήση τεχνικών αναλυτικής επεξεργασίας (online analytical processing –OLAP) και εξόρυξης γνώσης από δεδομένα (data mining), με απώτερο σκοπό την κατασκευή νέων εργαλείων και υπηρεσιών.

Όμως για την εκμετάλλευση τους απαιτείται η ανάπτυξη και χρήση ειδικευμένων αλγορίθμων. Με το γεγονός αυτό σαν αφετηρία τα τελευταία χρόνια η έρευνα έχει στραφεί στην εξερεύνηση των δυνατοτήτων που προσφέρει ο εμπλουτισμός των δεδομένων με σημασιολογική πληροφορία, θέμα που αποτελεί και το πλαίσιο της παρούσας εργασίας. Η εργασία βασίζεται πάνω στο σύστημα HERMES MOD (Níkos Pelekis, 2011), αξιοποιώντας την υποδομή που προσφέρει σε σημασιολογικά επαυξημένα δεδομένα κίνησης. Αντικείμενο της είναι η διερεύνηση θεμάτων που έχουν να κάνουν με την διαχείριση τέτοιων δεδομένων, και χωρίς να μένει εκεί προτείνει και επεκτάσεις για το HERMES, οι οποίες αφορούν τόσο στην αποτελεσματικότερη διαχείριση όσο και στην αξιοποίηση της πληροφορίας.

Η διάρθρωση του τεύχους έχει ως εξής:

Το κεφάλαιο 2 περιλαμβάνει το θεωρητικό υπόβαθρο της εργασίας, όπου γίνεται μία επισκόπηση της έννοιας των τροχιών και των σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων. Στη συνέχεια μέσα από τα παρουσιαζόμενα μοντέλα αναπαράστασης τους, εισάγεται η διάκριση των περιοχών α) σε βραδείας (ή και καθόλου) κίνησης, δηλώνονται σαν STOPS και σε β) ύπαρξης κίνησης, δηλώνονται σαν MOVES και κατ' επέκταση η έννοια των επεισοδίων. Στην ενότητα αυτή επίσης περιγράφεται και ο αλγόριθμος T-OPTICS, ο οποίος χρησιμοποιείται στην ανασύσταση των σημασιολογικών τροχιών καθώς και προτεινόμενες προσεγγίσεις ως προς τον σημασιολογικό εμπλουτισμό τους. Ακολούθως τίθενται θέματα διαχείρισης και επεξεργασίας αυτών και περιγράφεται το σύστημα HERMES MOD πάνω στο οποίο βασίζεται η εργασία.

Στο κεφάλαιο 3 ορίζονται τα κριτήρια απόδοσης και οι προσδοκίες χρήσης ενός συστήματος με αντικείμενο τις σημασιολογικά επαυξημένες τροχιές, που στην περίπτωση μας είναι το HERMES, καθώς και των επεκτάσεων του, ενώ μέσα από παραδείγματα γίνονται κατανοητά τα πλεονεκτήματα αυτής της προσέγγισης.

Στο 4^ο κεφάλαιο μελετάται η υπάρχουσα υποδομή και επεξηγείται η δομή της αποθήκης δεδομένων που προκύπτει. Ακόμη εμφανίζονται οι προτεινόμενες επεκτάσεις του συστήματος και οι σχετιζόμενοι αλγόριθμοι αυτών. Οι επεκτάσεις αφορούν θέματα διαχείρισης και επεξεργασίας τόσο κατά την διαδικασία της ανασύστασης όσο και κατά την διαδικασία της εκμετάλλευσης των δεδομένων για την παραγωγή γνώσης.

Στο 5^ο κεφάλαιο γίνεται η πρακτική εφαρμογή των παραπάνω σε δοθέν σύνολο δεδομένων. Παρουσιάζονται ποικίλα σημασιολογικά και χωροχρονικά ερωτήματα για την εξαγωγή αποτελεσμάτων με την εκτεταμένη χρήση τελεστών που το σύστημα HERMES διαθέτει. Επίσης, εξετάζονται και οι δυνατότητες των νέων επεκτάσεων με την εφαρμογή τους σε πραγματικά δεδομένα.

Η εργασία ολοκληρώνεται με το 6^ο κεφάλαιο στο οποίο γίνεται αποτίμηση της αποτελεσματικότητας του HERMES και των επεκτάσεων που προτείνουμε, εξετάζεται η κάλυψη ή μη των διατυπωθέντων προσδοκιών και στόχων και κλείνει με τα τελικά συμπεράσματα μας.

2 Θεωρητικό Υπόβαθρο

2.1 Τροχιές κινουμένων αντικειμένων (trajectories)

Ως χωροχρονικό φαινόμενο ορίζεται το φαινόμενο κατά το οποίο η θέση ή/και η έκταση μιας χωρικής οντότητας μεταβάλλεται με τον χρόνο. Κατά συνέπεια τα χωροχρονικά δεδομένα είναι δεδομένα που συνδυάζουν την χωρική πληροφορία με τη διάσταση του χρόνου και παράγονται από τη συστηματική καταγραφή των χωροχρονικών φαινομένων.

Απόρροια των παραπάνω είναι ότι η τροχιά T μιας οντότητας μπορεί να οριστεί ως η ακολουθία όλων των χωροχρονικών δεδομένων που παράγονται κατά την κίνησης της. Με την παραδοχή τώρα ότι η οντότητα είναι σημείο, η τροχιά του T , είναι η ακολουθία όλων των σημείων της θέσης του στο χώρο για κάθε χρονική στιγμή. Είναι δηλαδή μια συνεχόμενη αντιστοίχιση π.χ. για το δισδιάστατο επίπεδο από το $I \subseteq \mathbb{R} \rightarrow \mathbb{R}^2 (2D)$, αφού ορίζεται για κάθε χρονική στιγμή και ο χρόνος είναι άπειρο σύνολο.

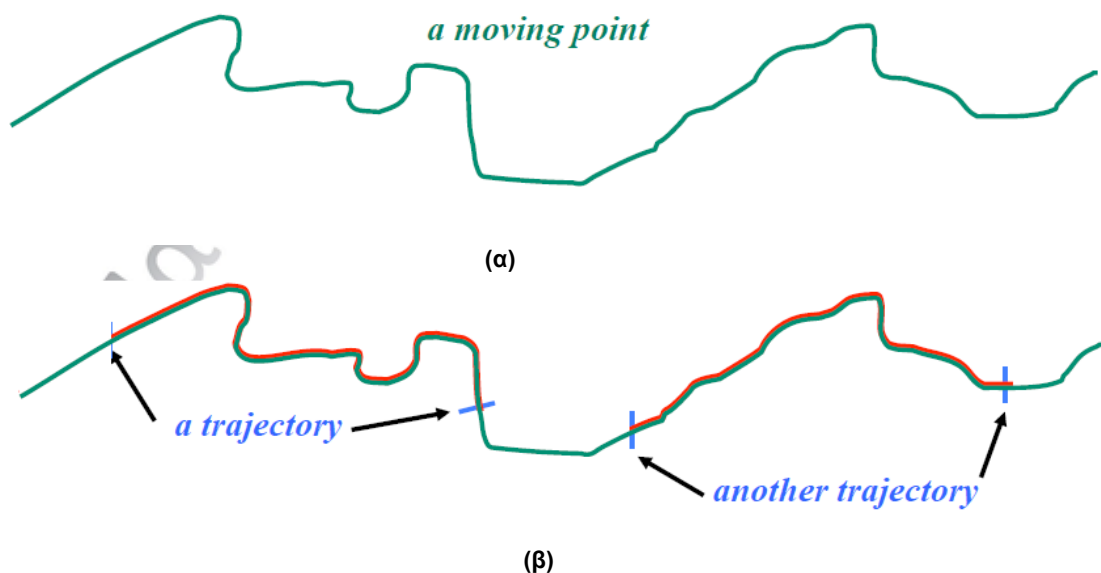
$$I \subseteq \mathbb{R} \rightarrow \mathbb{R}^2: t \rightarrow a(t) = (a_x(t), a_y(t))$$

$$\text{οπότε } T = (a_x(t), a_y(t), t) | t \in I \subset \mathbb{R}^2 \times \mathbb{R}$$

Βέβαια αν και η τροχιά είναι συνεχές φαινόμενο, οι διάφορες συσκευές και τεχνολογίες λειτουργούν διακριτά κατά την συλλογή δεδομένων, βασιζόμενες σε ποικίλα κριτήρια όπως είναι: ο χρονοπρογραμματισμός, η μεταβολή της τοποθεσίας, η ύπαρξη κάποιου ερεθίσματος κ.α. και εμπεριέχουν εγγενείς αδυναμίες, καθώς μπορούν να διαχειριστούν πληροφορία πεπερασμένου μεγέθους. Κρίνεται αναγκαία λοιπόν, η περιγραφή της τροχιάς με τη χρήση κάποιου διακριτού μοντέλου. Στην περίπτωση του δισδιάστατου επιπέδου ο καθορισμός της τρέχουσας θέσης περιγράφεται μετά από δειγματοληψία με μια τριπλέτα της μορφής (x, y, t) και η τροχιά T από μια πλειάδα τέτοιων τριπλετών.

$$T = ((x_1, y_1, t_1), (x_2, y_2, t_2), \dots, (x_n, y_n, t_n))$$

με $x_i, y_i, t_i \in \mathbb{R}$ και $t_1 < t_2 < \dots < t_n$.



Εικόνα 1: (α) Απεικόνιση της τροχιάς σαν συνεχές φαινόμενο και (β) απεικόνιση της σαν μια πεπερασμένη ακολουθία χρονοσημασμένων σημείων

Καθώς η τροχιά πλέον περιγράφεται από πλειάδες είναι αναγκαία μια διαδικασία ανασύστασης της. Η διαδικασία αυτή ονομάζεται *παρεμβολή* και σκοπός της είναι να συνδέσει με ικανοποιητικό τρόπο τα υπάρχοντα χρονοσημασμένα σημεία, ώστε η τελική καμπύλη κίνησης να προσεγγίζει την πραγματική. Ταυτόχρονα όμως θα πρέπει να αντιμετωπίζει τα προβλήματα ύπαρξης απεριόριστου αριθμού τροχιών που μπορεί να περνούν από τα δείγματα, ύπαρξης σφαλμάτων είτε κατά την συλλογή των δειγμάτων είτε κατά την εκτέλεση της διαδικασίας, αποδοτικότητας αλλά και αξιοπιστίας. Βλέπουμε, λοιπόν, ότι η ανασύσταση της τροχιάς δεν ορίζεται μονοσήμαντα και γι αυτό έχουν προταθεί διάφορες προσεγγίσεις ως προς το θέμα αυτό.

Μια από τις πιο διαδομένες τεχνικές είναι η *γραμμική παρεμβολή*, η οποία θεωρεί ότι μεταξύ δύο καταγεγραμμένων θέσεων μίας τροχιάς το αντικείμενο κινείται σε ευθεία γραμμή με σταθερή ταχύτητα. Η μέθοδος παρουσιάζει μικρές υπολογιστικές απαιτήσεις και για τον λόγο αυτό χρησιμοποιείται συχνά, όμως στην περίπτωση που η κίνηση είναι άτακτη και η δειγματοληψία αραιή παράγει αποτελέσματα με μικρή ακρίβεια. Στον αντίποδα είναι η μέθοδος των *πολύωνυμων splines* που μπορούν να χρησιμοποιηθούν για την καλύτερη εκτίμηση της θέσης του αντικειμένου σε χρονικές στιγμές για τις οποίες δεν υπάρχει πληροφορία. Αν και δίνει πολύ ικανοποιητικά αποτελέσματα, η πολυπλοκότητα της είναι αρκετά αυξημένη, οπότε προτιμάται συνήθως σε ειδικές περιπτώσεις.

Στη γραμμική παρεμβολή τα καμπυλόγραμμα τμήματα της τροχιάς έχουν αντικατασταθεί από ευθύγραμμα τμήματα και η τροχιά προσεγγίζεται από μια τεθλασμένη πολυγραμμή.

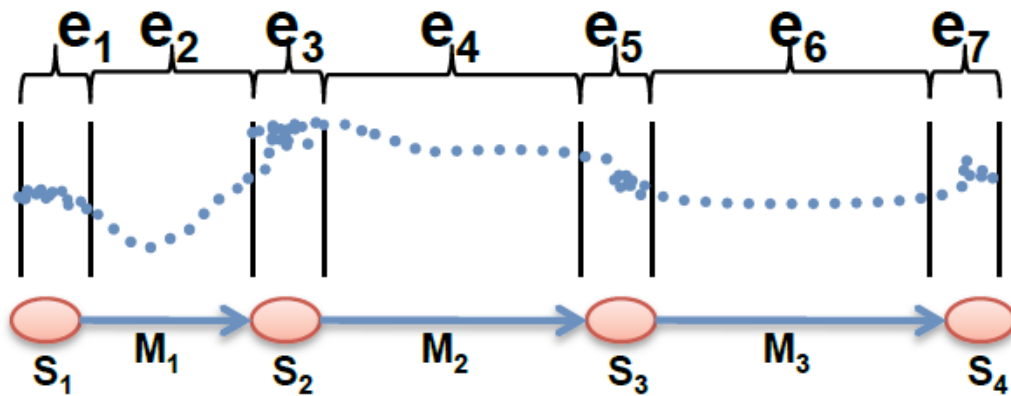
2.2 Σημασιολογικά επαυξημένες τροχιές (semantic trajectories)

2.2.1 Περιγραφή

Έχοντας δει παραπάνω την έννοια την τροχιών είναι εύκολο να διαπιστώσουμε ότι καλύπτουν ένα μέρος των απαιτήσεων που παρουσιάζονται κατά την ανάλυση δεδομένων κίνησης. Κι αυτό γιατί απουσιάζει το τμήμα της σημασιολογικής πληροφορίας της τροχιάς, όπως π.χ. ονόματα σημείων ενδιαφέροντος, εκδηλώσεις, σκοπός κίνησης κ.α. (application domain knowledge). Η διαδικασία λοιπόν κατά την οποία προστίθεται η πληροφορία αυτή στις τροχιές κινούμενων αντικειμένων ονομάζεται διαδικασία σημασιολογικού εμπλουτισμού και μπορεί να αφορά συνολικά την τροχιά ή τμήματα αυτής π.χ. η καταγραφή του σκοπού ταξιδιού ενός ατόμου (επαγγελματικοί λόγοι ή λόγοι αναψυχής) αφορά συνολικά στην τροχιά, ενώ η καταγραφή των μέσων μεταφοράς που χρησιμοποιεί σε διάφορα σημεία που αφορούν τμήματα αυτής.

Μια από τις μεθόδους που έχουν προταθεί είναι η πληροφορία αυτή να προστίθεται σαν τιμή κάποιας ιδιότητας π.χ. η ιδιότητα *ΜεταφορικόΜέσο* έχει τιμή "*λεωφορείο*". Η τιμή της εκάστοτε ιδιότητας μπορεί να προκύψει είτε μέσα από παρατηρήσεις, είτε μετά από υπολογισμό της από τα ακατέργαστα δεδομένα, είτε να εξαχθεί από σχετιζόμενα στοιχεία.

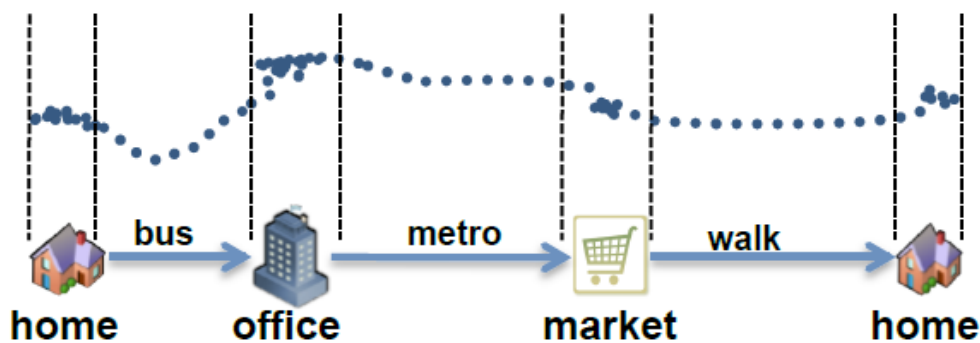
Επίσης ένας άλλος τρόπος προσθήκης γνώσης στις τροχιές γίνεται με τον λογικό διαχωρισμό τους σε χρονικά μη επικαλυπτόμενα τμήματα, τα οποία ονομάζονται *επεισόδια*. Τα επεισόδια ορίζονται ως μια υποακολουθία με ομοιογενή χαρακτηριστικά όπως χωροχρονικές συντεταγμένες των σημείων, σημασιολογικά δεδομένα και/ή χωροχρονικές σχέσεις με αντικείμενα τα οποία περιέχουν σημασιολογική πληροφορία (SPACCAPIETRA, 2011) και (Zhixian Yan N. G., 2011). Με την προσέγγιση αυτή παράγεται μια πιο κατατοπιστική αναπαράσταση της με την μορφή διατεταγμένης ακολουθίας επεισοδίων.



Εικόνα 2: Απεικόνιση μιας τροχιάς με την μορφή ακολουθίας επεισοδίων.

Η αναπαράσταση με επεισόδια εισάγει το πλεονέκτημα της χρήσης αλγορίθμων για την αυτόματη παραγωγή τους με βάση χαρακτηριστικά όπως είναι η ταχύτητα, η επιτάχυνση, ο προσανατολισμός, η πυκνότητα κ.α. αλλά και της συμπίεσης της πληροφορίας, καθώς δεν είναι απαραίτητος πλέον ο σημασιολογικός εμπλουτισμός κάθε εγγραφής ξεχωριστά αλλά αρκεί ο εμπλουτισμός των επεισοδίων με όμοια χαρακτηριστικά. Η τροχιά η οποία πλέον εμφανίζεται σαν ακολουθία επεισοδίων μπορούμε να πούμε ότι είναι *εννοιολογικά επαυξημένη*, ενώ ένας συνηθής διαχωρισμός αφορά την ύπαρξη ή όχι της κίνησης. Αυτό μας οδηγεί σε δύο τύπους επεισοδίων MOVES ή STOPS.

Επέκταση της εννοιολογικά επαυξημένης τροχιάς είναι η *σημασιολογικά επαυξημένη τροχιά*, η οποία ορίζεται ως αυτή η οποία έχει εμπλουτιστεί με σημασιολογική πληροφορία - σχολιασμούς (annotations) και /ή με ένα ή περισσότερα τμήματα-επεισόδια (Zhixian Yan N. G., 2011). Ο σχολιασμός που προστίθεται μπορεί να αφορά είτε συνολικά στην τροχιά, είτε στα επεισόδια αυτής, ενώ όπως είπαμε και πιο πάνω αποφεύγεται η προσθήκη πληροφορίας σε διακριτά σημεία-εγγραφές της, καθώς κάτι τέτοιο οδηγεί στην ύπαρξη πλεονασμού και επαναλήψεων.



Εικόνα 3: Απεικόνιση σημασιολογικά επαυξημένης τροχιάς

2.2.2 Μοντέλα αναπαράστασης-επεξεργασίας σημασιολογικά επαυξημένων τροχιών

Για την μοντελοποίηση της σημασιολογικά επαυξημένης τροχιάς έχουν γίνει ποικίλες μελέτες και προτάσεις. Μια από τις πρώτες προσεγγίσεις είναι αυτή του Spaccapietra (Spaccapietra S, 2008), κατά την οποία η σημασιολογική πληροφορία σχετίζεται μόνο με τα μέρη όπου το κινούμενο αντικείμενο σταμάτησε ή κινήθηκε. Στη περίπτωση αυτή όμως εμπεριέχεται ένας γηγενής περιορισμός, καθώς μπορεί να απεικονιστεί πληροφορία μόνο στην τροχιά και όχι σε σημεία της τροχιάς. Με μια νεότερη εργασία ο

Yan (Zhixian Yan D. C., 2011) δίνει μια πιο πλήρη περιγραφή των εννοιών της σημασιολογικά επαυξημένης τροχιάς, χωρίς να έχει αυστηρά την μορφή ορισμού. Στην μελέτη του προτείνεται ένα πλαίσιο εργασίας με τρία επίπεδα: το επίπεδο των δεδομένων, όπου η τροχιά ορίζεται σαν διατεταγμένη ακολουθία σημείων με αρχή και τέλος, το εννοιολογικό και σημασιολογικό επίπεδο, όπου γίνεται η κατάτμηση της τροχιάς σε υποακολουθίες - επεισόδια, και ο σημασιολογικός εμπλουτισμός της τροχιάς αντίστοιχα.

Επίσης στην πρόταση του Yan (Zhixian Yan N. G., 2011) ορίζονται οι βασικές έννοιες και η ορολογία για την υποστήριξη της ανάλυσης των τροχιών, των συμπεριφορών τους καθώς και των βημάτων για την απευθείας μετατροπή των αρχικών ακατέργαστων δεδομένων σε σημασιολογικά επαυξημένα. Σε μια άλλη εργασία, αυτή της Bogorny's (Vania Bogorny C. R., 2013) εισάγεται ένα γενικότερο σημασιολογικό μοντέλο που αντιμετωπίζει τα επεισόδια STOP και MOVES σαν υποπεριπτώσεις ομοειδών χαρακτηριστικών των εγγραφών, ενώ ορίζεται και οι έννοια της σημασιολογικής υποτροχιάς. Με την πρόταση αυτή η σημασιολογική πληροφορία μπορεί να μοντελοποιηθεί πλήρως αλλά και να υποστηρίξει πιο αποτελεσματικά ερωτήματα.

2.2.3 Σημασιολογικός εμπλουτισμός της τροχιάς

Η διαδικασία του σημασιολογικού εμπλουτισμού μπορούμε να πούμε ότι χωρίζεται σε δύο στάδια, στην κατάτμηση της τροχιάς σε επεισόδια και στον εμπλουτισμό της με σημασιολογική πληροφορία.

Κατάτμηση τροχιάς σε επεισόδια

Όπως είδαμε για την εννοιολογική αναπαράσταση της τροχιάς είναι αναγκαία η τμηματοποίηση της σε επεισόδια. Η πιο συνήθης διαδικασία αφορά στη διάκριση των περιοχών α) σε βραδείας (ή και καθόλου) κίνησης, δηλώνονται σαν STOPS και σε β) ύπαρξης κίνησης, δηλώνονται σαν MOVES. Εδώ οι μέθοδοι που προτείνονται αφορούν κυρίως των εντοπισμό των STOPS και μπορούν να διακριθούν σε δύο κατηγορίες: τις *στατικές* και τις *δυναμικές*.

Στις στατικές μεθόδους δίνεται η γεωγραφική τοποθεσία σημαντικών θέσεων, Σημεία Ενδιαφέροντος (Point Of Interests (POI), Περιοχές/Γραμμές Ενδιαφέροντος (Region/Line of Interests - ROI/LOI). Ως προς αυτά εξετάζεται πρώτιστα τυχόν χωρική διασταύρωση της τροχιάς, ενώ μεταγενέστερα εισέρχονται πρόσθετοι περιορισμοί (π.χ. χρονική διάρκεια), ώστε να φιλτραριστούν καλύτερα τα αποτελέσματα. Γίνεται φανερό ότι η αποτελεσματικότητα τέτοιων μεθόδων εξαρτάται από την αξιοπιστία των POI/ROI/LOI δεδομένων, γεγονός που τις περιορίζει σημαντικά.

Στον αντίποδα, βρίσκονται οι δυναμικές μέθοδοι όπου απουσιάζει η a priori γνώση και για τον εντοπισμό των STOPS βασίζονται σε χωροχρονικά χαρακτηριστικά της τροχιάς, όπως η ταχύτητα, η επιτάχυνση, η κατεύθυνση κ.α. Σαν παράδειγμα μπορούμε να αναφέρουμε την εργασία του Agamennoni (G. Agamennoni, 2009) στην οποία παρουσιάζεται μια συνάρτηση αξιολόγησης για την δημιουργία συστάδων που αντιστοιχούν σε STOPS με βάση ένα σταθερό κατώφλι ταχύτητας. Επίσης σε μια άλλη εργασία αυτή του Yan (Z. Yan, 2010) γίνεται χρήση ενός δυναμικού κατωφλίου ταχύτητας (Δ_{speed}) για την αναγνώριση των STOPS. Με αυτό οι συνεχόμενες εγγραφές με ταχύτητα κάτω από Δ_{speed} λαμβάνονται σαν STOPS ενώ διαφορετικά σαν MOVES.

Εκτός όμως από την χρήση κατωφλίου για την παραγωγή αποτελεσμάτων, σε άλλες περιπτώσεις γίνεται εξέταση της πυκνότητας με την χρήση γνωστών αλγορίθμων συσταδοποίησης. Ο Ashbrook (Starner, 2003) για παράδειγμα χρησιμοποιεί τον αλγόριθμο *k-means* για να ανακαλύψει θέσεις με ενδιαφέρον από ιστορικά GPS δεδομένα, ενώ ο Zimmermann (M. Zimmermann, 2009) επεκτείνει τον OPTICS, ώστε να ανακαλύψει STOPS. Η παραλαγή αυτή ονομάζεται T-OPTICS και επειδή είναι ενσωματωμένη στο σύστημα Sem-Hermes στο οποίο βασίστηκε η εν λόγω εργασία θα παρουσιαστεί πιο αναλυτικά παρακάτω.

Ο αλγόριθμος T-OPTICS

Ο T-OPTICS αποτελεί μια παραλλαγή του αλγορίθμου OPTICS, ο οποίος ανήκει στην κατηγορία των αλγορίθμων συσταδοποίησης με βάση την πυκνότητα και ο οποίος με την σειρά του αποτελεί επέκταση του DBSCAN. Ο σκοπός ορισμού του αφορά την εύρεση περιοχών κίνησης και στασιμότητας σε Διαχείριση σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων

δεδομένα τροχιάς και γι αυτό εισάγονται σε αυτόν τόσο η χωρική όσο και η χρονική διάσταση. Έτσι κατά αντιστοιχία των αλγορίθμων πυκνότητας, όταν ορίζεται η έννοια της "γειτονίας" λαμβάνεται υπόψη εκτός από την χωρική και η χρονική απόσταση. Με τον τρόπο αυτό εξασφαλίζεται το ότι δεν αποτελούν μέρος της ίδιας γειτονίας τα σημεία που αν και είναι κοντά χωρικά, απέχουν χρονικά αλλά και η δυνατότητα εξέτασης συνεχόμενων και όχι τυχαίων σημείων.

Στη συνέχεια ο αλγόριθμος βασιζόμενος σε αυτή την τροποποιημένη προσέγγιση της γειτονίας και με την παραδοχή ότι οι συστάδες αναπαριστούν STOPS χρησιμοποιεί τις τιμές προσεγγισιμότητας όλων των σημείων και καθορίζει μια συστάδα ως την ακολουθία σημείων ανάμεσα σε δύο μεγάλες τιμές προσεγγισιμότητας. Αυτά τα σημεία της τροχιάς είναι χρονικά εγγυή και χωρικά κοντά οπότε αντιστοιχούν σε μια περιοχή βραδείας (ή και καθόλου) κίνησης-STOP.

Σημασιολογικός Εμπλουτισμός

Η προσθήκη σχολιασμού στα επεισόδια μπορεί να αφορά στον ορισμό τους π.χ. STOP ή MOVE αλλά μπορεί να αφορά και σε πρόσθετη σχετιζόμενη με αυτά πληροφορία όπως π.χ. σκοπός της στάσης (ψώνια, ψυχαγωγία) ή τρόπος μετακίνησης (λεωφορείο, τρένο) αντίστοιχα. Σε ένα ανώτερο επίπεδο ο χαρακτηρισμός ολόκληρης της τροχιάς προϋποθέτει τη σύνθεση συνολικά της πληροφορίας που τυχόν την χαρακτηρίζει. Για παράδειγμα αν σε μια τροχιά το αρχικό και τελικό επεισόδιο είναι αντίστοιχα "ΣΠΙΤΙ", "ΕΡΓΑΣΙΑ", τότε θα μπορούσε να εμπλουτιστεί με την ετικέτα "ΣΠΙΤΙ ΠΡΟΣ ΕΡΓΑΣΙΑ". Αν και μερικές από τις προτεινόμενες προσεγγίσεις μοντελοποίησης αναφέρθηκαν στην παράγραφο 2.2.2. θα παραθέσουμε κάποιες βασικές έννοιες, οι οποίες σχετίζονται με το κομμάτι του σημασιολογικού εμπλουτισμού.

Σε πολλές μελέτες συναντάμε την έννοια της *συμπεριφοράς*, η οποία συντίθεται από ένα σύνολο χαρακτηριστικών, τα οποία προσδίδουν ένα ιδιαίτερο νόημα σε ένα κινούμενο αντικείμενο ή σύνολο κινούμενων αντικειμένων (Vania Bogorny C. R., 2013). Σε γενικές γραμμές η συμπεριφορά μιας τροχιάς ή ενός συνόλου τροχιών δεν μπορεί να γίνει κατανοητή ή να εξαχθεί με απλή εξέταση των δεδομένων ή εφαρμογή ερωτημάτων αλλά μέσα από τεχνικές εξόρυξης γνώσης. Μια τροχιά κατά την συνολική της διάρκεια μπορεί να εμφανίζει διαφορετικές συμπεριφορές στις υποτροχιές (επεισόδια) της.

Η συμπεριφορά μπορεί να διακριθεί σε *ατομική* (*single*) ή *συλλογική* (*collective*) ανάλογα με το αν εμφανίζεται σε μια μόνο τροχιά ή σε ένα σύνολο τροχιών. Η συλλογική συμπεριφορά μπορεί περαιτέρω να διακριθεί (Vania Bogorny C. R., 2013) σε γνωρίζουσα (*aware*) ή μη-γνωρίζουσα (*non-aware*). Η κύρια διαφορά μεταξύ αυτών των δύο είναι ότι στην πρώτη ένα τουλάχιστον κινούμενο αντικείμενο έχει επίγνωση των άλλων κινούμενων αντικειμένων και η κίνηση του εξαρτάται από την κίνηση τους. Με άλλα λόγια η τροχιά του επηρεάζεται από άλλες τροχιές και η συλλογική κίνηση που προκύπτει δεν είναι αποτέλεσμα κάποιου τυχαίου γεγονότος. Το αντίθετο συμβαίνει στην δεύτερη κατηγορία, όπου ένα σύνολο τροχιών έχει παρόμοια συμπεριφορά από σύμπτωση ή όταν δεν έχει πρόθεση να παραμείνει με τις άλλες τροχιές.

Επίσης, υπάρχει η *χωρο-χρονική συμπεριφορά* (*spatio-temporal behavior*), στην οποία οι ιδιότητες της τροχιάς εξαρτώνται μόνο από χωρικά ή χρονικά δεδομένα αλλά και η *σημασιολογική συμπεριφορά* (*semantic behavior*), τα χαρακτηριστικά της οποίας πηγάζουν από συνδυασμό σημασιολογικών δεδομένων και κάποιων χωρικών ή/και χρονικών δεδομένων.

2.3 Βάσεις κινούμενων αντικειμένων (MOD) - HERMES

Η έρευνα για τις βάσεις κινούμενων αντικειμένων λαμβάνει υπόψη την ανάγκη για την αναπαράσταση της κίνησης, προκειμένου να γίνεται δυνατή η εκτέλεση ειδικών ερωτημάτων και η ανάλυση τους. Μια σειρά από έγγραφα αντιμετωπίζουν θέματα όπως είναι η μοντελοποίηση των δεδομένων, η ανάπτυξη γλωσσών για την εκτέλεση ερωτημάτων και η υλοποίηση τους. Ιδιαίτερα ως προς το τελευταίο έχουν προταθεί δύο πρότυπες εργασίες, το SECONDO (R.H. Güting, 2004) και το HERMES (Nikos Pelekis, 2011). Το σύστημα HERMES πάνω στο οποίο βασίζεται και η εν λόγω διπλωματική, αποτελεί ένα ολοκληρωμένο πλαίσιο επέκτασης των βάσεων δεδομένων βασισμένο στην εργασία του Güting (R.H. Güting, 2000) για χωροχρονικά δεδομένα. Πιο συγκεκριμένα, για τον ορισμό του μοντέλου γίνεται χρήση χρονικών τύπων (*temporal types*), οι οποίοι εισήχθησαν μέσω της βιβλιοθήκης *TAU Temporal Literal Library (TAU-TLL)* (N.Pelekis, 2002), καθώς και χωρικών τύπων, οι οποίοι συναντώνται σε όλα τα σύγχρονα συστήματα βάσεων δεδομένων. Ο πυρήνας της αρχιτεκτονικής του συστήματος είναι ένα

Oracle Data Cartridge, που ονομάζεται HERMES Moving Data Cartridge (HEMRES-MDC) και με βάση το οποίο προσφέρονται πλήθος προηγμένων λειτουργιών ανασύστασης, διαχείρισης, επεξεργασίας και ανάλυσης δεδομένων κινούμενων αντικειμένων, καθώς και δεδομένων μεταβαλλόμενων στο χρόνο γεωμετριών (κύκλος, πολύγωνο, σημείο κλπ).

Πανεπιστήμιο Πειραιώς

3 Ανάλυση Απαιτήσεων

Οι γενικές απαιτήσεις της εργασίας αφορούσαν στην εξερεύνηση των δυνατοτήτων ενός προηγμένου συστήματος, όπως είναι το HERMES καθώς και την προσθήκη σε αυτό επεκτάσεων σχετικά με τη διαχείριση και επεξεργασία των σημασιολογικά επαυξημένων τροχιών. Έγινε χρήση των υπαρχουσών λειτουργιών του συστήματος και ανάπτυξη νέων με χρήση γλώσσας προγραμματισμού PL/SQL σε περιβάλλον Oracle 11G R2 Enterprise Edition και δυνατότητας Oracle Spatial.

3.1 Κριτήρια αξιολόγησης

Η εκμετάλλευση με αποτελεσματικό και αξιόπιστο τρόπο της πληροφορίας είναι το πρώτο κριτήριο που αφορά την εργασία αυτή. Ως προς αυτό το σκοπό εξετάζονται διαφορετικά σενάρια πάνω σε δοθέν σύνολο δεδομένων. Τα σενάρια αυτά αφορούν σε ερωτήματα χρονικού, χωρικού και χωροχρονικού εύρους καθώς και στην εφαρμογή διαφορών τελεστών για την καλύτερη εύρεση απαντήσεων με τη χρήση της σημασιολογικής πληροφορίας. Τέτοιου είδους ερωτήματα είναι π.χ. (α) “*Βρες τα άτομα που ακολουθούν την διαδρομή: σπίτι – δουλεία από τις 7:00πμ έως τις 10:00πμ*”, (β) “*Πόσοι άνθρωποι χρησιμοποιούν μέσα μαζικής μεταφοράς για να πάνε στην εργασία τους*” για τα οποία σε αντίθετη περίπτωση είναι αναγκαία η χρήση πιο προηγμένων τεχνικών, όπως είναι π.χ. οι μέθοδοι εξόρυξης γνώσης. Για να γίνουν καλύτερα κατανοητά τα πλεονεκτήματα που προκύπτουν ακολουθεί ένα παράδειγμα. Ας υποθέσουμε για το ερώτημα: “*Ποια είναι τα μέρη από τα οποία ένα άτομο πέρασε κατά την διάρκεια της τροχιάς του;*”, ότι για κάθε κινούμενο αντικείμενο υπάρχει μια μόνο τροχιά στην βάση δεδομένων και οι γεωγραφικές θέσεις που μας ενδιαφέρουν είναι δύο *γήπεδο* και *ξενοδοχείο*. Για την περίπτωση όπου απουσιάζει η σημασιολογική πληροφορία, εάν υποθέσουμε ότι το κινούμενο αντικείμενο *A* έχει την τροχιά (1), όπως αυτή φαίνεται στην Εικόνα 4 χωρίς ύπαρξη σημασιολογικής πληροφορίας, τότε για την απάντηση του ερωτήματος θα απαιτούνταν ένα SQL ερώτημα με δυο ενώσεις με χωρικά δεδομένα (spatial joins), όπως το παρακάτω:

```
SELECT h.NAME
FROM trajectory t, hotel h
WHERE t.ID = 'A' AND intersects (t.moving_point.geometry, h.geometry)
UNION
SELECT s.NAME
FROM trajectory t, stadium s
WHERE t.ID = 'A' AND intersects (t.moving_point.geometry,
s.geometry)
```

Η περίπτωση αυτή είναι απλή και ο τελεστής της (**intersects**) θα εξεταζόταν, έως ότου βρισκόταν τομή κάποιου σημείου της τροχιάς με τις δοθείσες θέσεις. Όταν όμως εισαχθεί και χρονικός περιορισμός (χωροχρονικός περιορισμός) τότε αναγκαία η συντακτική και υπολογιστική πολυπλοκότητα του SQL πρέπει να αυξηθεί. Π.χ ρωτώντας, “*Ποια είναι τα μέρη στα οποία ο A έμεινε για περισσότερο από 1 ώρα;*” η αναζήτηση δεν μπορεί να τερματιστεί, όταν το πρώτο σημείο της τροχιάς τέμνει τη δοθείσα θέση, αλλά όλα τα σημεία της τροχιάς που τέμνονται πρέπει να εξεταστούν, μέχρι το κινούμενο αντικείμενο να φύγει από τη θέση ενδιαφέροντος, ή να ικανοποιηθεί ο χρονικός περιορισμός.

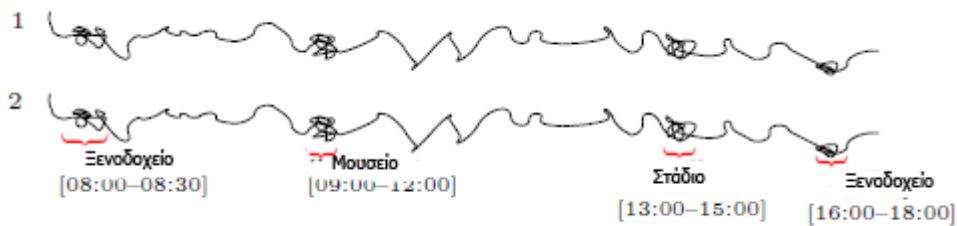
Σε αντιδιαστολή, κατά την χρήση σημασιολογικά επαυξημένης τροχιάς βλ. Εικόνα 4 (2) στην οποία εμπεριέχεται η πληροφορία των γεωγραφικών θέσεων, το SQL για το ίδιο παράδειγμα θα είχε ως εξής:

```
SELECT st.place
```

```
FROM semantic_trajectory st
```

```
WHERE t.ID = 'A'
```

Είναι φανερό ότι είναι συντακτικά κατά πολύ απλούστερο αλλά και δεν απαιτεί, καθώς πλέον ενσωματώνει την πληροφορία, τον υπολογισμό της συσχέτισης ολόκληρης της τροχιάς με κάθε γεωγραφικό σημείο για να απαντηθεί το ερώτημα. Δηλαδή σε αυτή την περίπτωση η σημασιολογική και χωροχρονική πληροφορία είναι ήδη συνδεδεμένες και η απάντηση τέτοιων ερωτημάτων μπορεί να γίνει με ερωτήματα προβολής, επιλογής. Κατά την περίπτωση όμως όπου απουσιάζει η σημασιολογική πληροφορία (πρώτο παράδειγμα) είναι απαραίτητος αυτός ο υπολογισμός για κάθε ερώτημα, το οποίο αφορά είτε το ίδιο, είτε διαφορετικό γεωγραφικό σημείο.



Εικόνα 4: (1) μια απλή τροχιά και (2) μια απλή σημασιολογικά επαυξημένη τροχιά.

Ως δεύτερο κριτήριο ορίζεται η ανάπτυξη επεκτάσεων στο υπάρχον σύστημα. Επιλέχθηκαν η δυνατότητα οπτικοποίησης μιας σημασιολογικής τροχιάς, με την χρήση της Keyhole Markup Language (KML), καθώς και αλγόριθμοι για την εύρεση α) των K κοντινότερων σημείων ενδιαφέροντος από το κεντροειδές του ελαχίστου περιβάλλοντος ορθογωνίου (MBR) ενός επεισοδίου και β) των πιθανοτήτων κάθε σημείου να είναι το σημείο στο οποίο πραγματοποιήθηκε η δραστηριότητα ενός επεισοδίου. Και οι δυο αυτοί αλγόριθμοι συνεισφέρουν στη λειτουργία σύνθεσης μιας σημασιολογικά επαυξημένης τροχιάς. Επίσης, περιλήφθηκαν διαδικασίες δημιουργίας μητρών προέλευσης - προορισμού (origin-destination matrix) τόσο από τις σημασιολογικές τροχιές του διαθέσιμου συνόλου δεδομένων όσο και από την επαγόμενη πληροφορία της αντίστοιχης αποθήκης δεδομένων.

3.2 Προσδοκίες χρήσης

Η εκμετάλλευση των τελεστών του HERMES για σημασιολογικά επαυξημένες τροχιές καθώς και η αποτελεσματικότητα των νεοεισαχθέντων διαδικασιών αποτελούν τις προσδοκίες χρήσης της παρούσας εργασίας. Η αναμενόμενη κατάληξη λοιπόν πρέπει να περιλαμβάνει:

- Την αποδοτική εφαρμογή των δυνατοτήτων του HERMES σε διαφορετικά σενάρια εκμετάλλευσης των δεδομένων (χωρικών, χρονικών, χωροχρονικών, σημασιολογικών περιορισμών).
- Την χρήση των νέων τελεστών είτε μεμονωμένα είτε σε συνεργασία με τους υπάρχοντες για την εξαγωγή επιτυχών αποτελεσμάτων.
- Την αναπαράσταση τους όπου είναι δυνατόν.

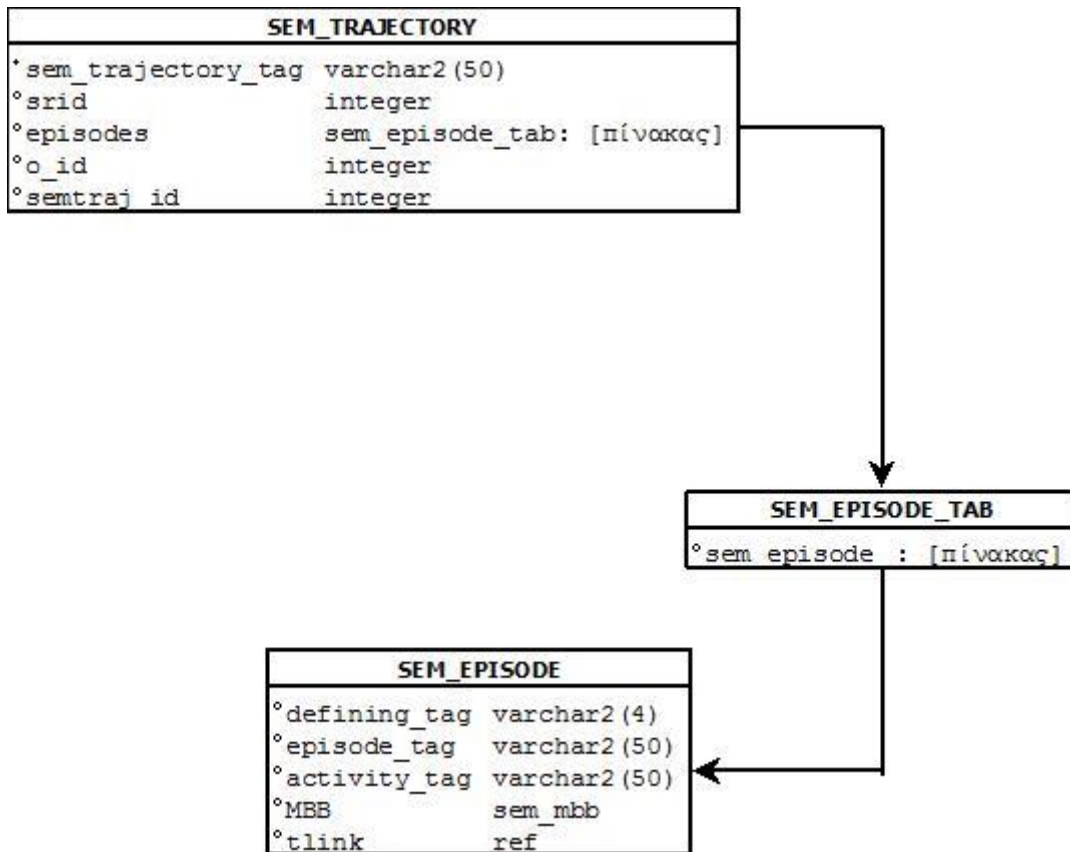
4 Σχεδιασμός

4.1 Υπάρχουσα δομή

4.1.1 Σημασιολογικά επαυξημένες τροχιές στο HERMES MOD

Στην υπάρχουσα μοντελοποίηση ο κύριος τύπος που χρησιμοποιείται είναι αυτός της σημασιολογικά επαυξημένης τροχιάς (SEM_TRAJECTORY). Ο τύπος αυτός ορίζεται από πλειάδα της μορφής (*o-id*, *semtraj-id*, T_{sem} , *sem_traj_tag*, *srid*), όπου το *o-id* (*semtraj-id*) είναι το αναγνωριστικό του κινούμενου αντικειμένου (και της σημασιολογικής τροχιάς του κινούμενου αντικειμένου αντίστοιχα), το T_{sem} είναι μια ακολουθία επεισοδίων τα οποία ανήκουν στην ίδια τροχιά τ και τα οποία δεν επικαλύπτονται χρονικά, δηλαδή ισχύει: $e_i[t_{end}] \leq e_{i+1}[t_{start}]$. Επίσης, τα *sem_traj_tag*, *srid* αφορούν στην ετικέτα που χαρακτηρίζει συνολικά την τροχιά και το γεωγραφικό σύστημα συντεταγμένων στο οποίο έχει γίνει η δειγματοληψία της.

Ο τύπος για τα επεισόδια δηλώνεται από πλειάδα της μορφής (*defining_tag*, *episode_tag*, *activity_tag*, *MBB*, *T-link*), όπου η πρώτη ιδιότητα είναι μια ένδειξη για STOP/MOVE, ενώ η δεύτερη και τρίτη αφορούν τον σημασιολογικό χαρακτηρισμό της θέσης όπου έλαβε χώρα το επεισόδιο και τον σκοπό του αντίστοιχα. Το MBB (τύπος SEM_MBB στο HERMES) είναι μια τριάδα της μορφής (MBR, t_{start} , t_{end}) που αντιστοιχεί σε μια χωροχρονική 3D προσέγγιση της υπο-τροχιάς τ' (περιγράφεται παρακάτω) του επεισοδίου. Για την ακρίβεια το MBR (Minimum Bounding Rectangle) είναι το δισδιάστατο περιβάλλον ορθογώνιο της χωρικής προβολής της τ' , και τα $[t_{start}, t_{end}]$ είναι η μονοδιάστατη χρονική προβολή της. Ακόμη υπάρχει για τη σύνδεση του επεισοδίου με τα χωροχρονικά δεδομένα της μορφής (x,y,t), ο δείκτης *T-link* (TLINK). Το σύνολο αυτών των δεδομένων κάθε επεισοδίου ονομάζεται υπο-τροχιά (sub-trajectory), ώστε να γίνεται εύκολα αντιληπτό ότι κάθε τροχιά αποτελείται από υπο-τροχιές.



Εικόνα 5: Σχηματική απεικόνιση μέρους των αντικειμένων της επέκτασης του HERMES.

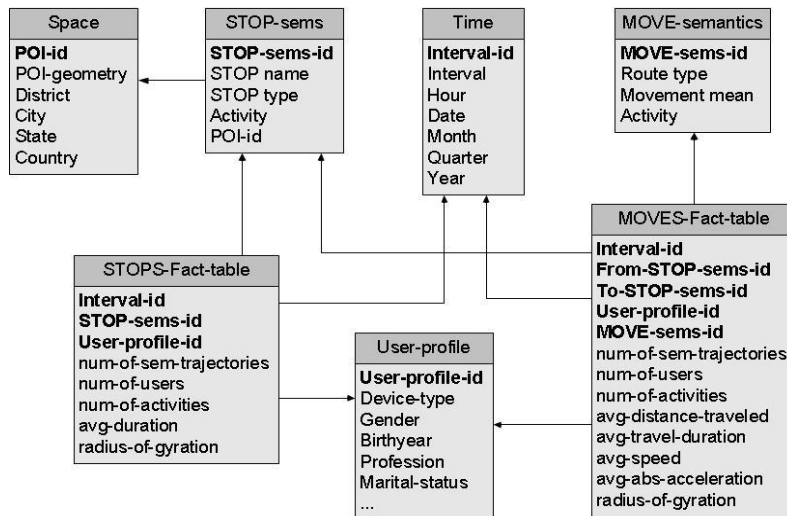
4.1.2 Αποθήκη δεδομένων σημασιολογικά επαυξημένων τροχιών στο HERMES MOD

Μέρος της υπάρχουσας δομής του HERMES MOD αποτελεί και η ύπαρξη αποθήκης δεδομένων σχήματος αστερισμού για την υποστήριξη της επεξεργασμένης πληροφορίας των σημασιολογικών τροχιών. Καθώς οι κύριες οντότητες μιας τροχιάς είναι τα MOVE και STOP επεισόδια, αυτά επιλέγονται και σαν πίνακες γεγονότων (MOVES-fact-table και STOPS-fact-table). Για κάθε έναν από αυτούς, κατά την τροφοδότηση τους μέσω της ETL διαδικασίας ορίζονται βασικά κελιά (*base cells – bc*), ώστε να μπορεί να γίνει ο διαχωρισμός του πολυδιάστατου χώρου των τροχιών και κατά συνέπεια και ο υπολογισμός των μετρήσιμων μεγεθών. Για την ακρίβεια κατά την τροφοδοσία του *stops-fact-table* ορίζουμε το bc_{stop} , το οποίο συντίθεται από τρεις διαστάσεις (χρόνος, προφίλ-χρήστη και σημασιολογική πληροφορία του STOP επεισοδίου), ενώ αντίστοιχα στην περίπτωση του *moves-fact-table* ορίζουμε bc_{moves} το οποίο αποτελείται από πέντε διαστάσεις (χρόνος, προφίλ-χρήστη, σημασιολογική πληροφορία του MOVE επεισοδίου, σημασιολογική πληροφορία του STOP επεισοδίου που προηγήθηκε του MOVE και σημασιολογική πληροφορία του STOP επεισοδίου που ακολούθησε του MOVE), στην συνέχεια βασιζόμενοι στα κελιά αυτά υπολογίζουμε τις τιμές των μεγεθών για τους αντίστοιχους πίνακες γεγονότων. Για παράδειγμα για το μέγεθος *num-of-sem-trajectories* προσμετράται ο αριθμός των διακριτών τροχιών που βρίσκονται σε ένα κελί, ενώ το *num-of-users* καθορίζεται από τον αριθμό των διαφορετικών κινουμένων αντικειμένων ενός κελιού κ.ο.κ.

Ολοκληρωμένο το σχήμα αποτελείται από πέντε πίνακες διαστάσεων: χώρος, χρόνος, προφίλ χρηστών, *move-semantic*, *stop-semantic*, και τους δύο πίνακες γεγονότων. Οι στήλες όλων των πινάκων δείχνονται στην Εικόνα 6 και είναι αυτοπεριγραφόμενες.

Πιο συγκεκριμένα ο πίνακας *Time* περιέχει παραγόμενη πληροφορία από την κατάτμηση της χρονικής περιόδου του συνόλου των τροχιών. Υπάρχει η δυνατότητα η κατάτμηση να γίνει με χρονικό βήμα επιλεγόμενο από τον χρήστη και μπορεί να είναι σε μήνες, μέρες κλπ. Ο πίνακας *Space* αφορά στα Διαχείριση σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων

σημεία ενδιαφέροντος καθώς και ιδιότητες αυτών και ο *User-profile* πληροφορία ως προς τα χαρακτηριστικά των χρηστών.



Εικόνα 6: Σχηματική απεικόνιση της αποθήκης δεδομένων για σημασιολογικά επαυξημένες τροχιές στο HERMES.

4.2 Σχεδιασμός ερωτημάτων

Ο σχεδιασμός έγινε με τέτοιο τρόπο, ώστε να γίνει εκμετάλλευση και ανάδειξη των δυνατοτήτων του HERMES MOD ως προς τις σημασιολογικά επαυξημένες τροχιές. Για τον λόγο αυτό περιλαμβάνονται απλά ερωτήματα με την χρήση τελεστών πάνω σε μεμονωμένες τροχιές, χωρικά, χρονικά, χωροχρονικά ερωτήματα επιλογής / προβολής με χωρικούς, χρονικούς, χωροχρονικούς και σημασιολογικούς περιορισμούς, με τη χρήση ευρετηρίων (STB-tree) ή όχι καθώς και πλήθος παραλλαγών αυτών.

4.3 Επεκτάσεις

Αναλυτικά οι επεκτάσεις του συστήματος έχουν ως εξής:

4.3.1 Απεικόνιση σημασιολογικά επαυξημένης τροχιάς.

Η προσθήκη αυτή μας βοηθά να απεικονίσουμε ολοκληρωμένα μια σημασιολογική τροχιά (είσοδος) με την παραγωγή αρχείων KML και την χρήση Google Maps. Κατά την απεικόνιση της μπορούν να δειχθούν κατ' επιλογή τα ελάχιστα περιβάλλοντα ορθογώνια των επεισοδίων της, η σημασιολογική πληροφορία αυτών καθώς και οι ανάλογες υπο-τροχιές τους.

4.3.2 Εύρεση K πλησιέστερων σημείων ενδιαφέροντος ενός επεισοδίου.

Με την επέκταση αυτή βρίσκουμε τα K (είσοδος) κοντινότερα σημεία ενδιαφέροντος (Points of Interest) από το κεντροειδές του ελάχιστου περιβάλλοντος ορθογώνιου (MBR, είσοδος) ενός επεισοδίου. Τα σημεία ενδιαφέροντος πρέπει να βρίσκονται εντός ή να εφάπτονται του MBR. Ανάλογα με τον γεωμετρικό περιορισμό, ο οποίος καθορίζεται από το MBR του επεισοδίου δύναται να επιστρέφονται λιγότερα από K σημεία, εφόσον βέβαια αυτά υπάρχουν. Επίσης δίνεται και η δυνατότητα επιλογής απεικόνισης ή όχι των αποτελεσμάτων.

4.3.3 Εύρεση πιθανοτήτων για τα K πλησιέστερα σημεία ενδιαφέροντος ενός επεισοδίου.

Με τη διαδικασία αυτή βρίσκουμε εκείνα τα σημεία ενδιαφέροντος τα οποία βρίσκονται μέσα στο ελάχιστο περιβάλλον ορθογώνιο (MBR, είσοδος) ενός επεισοδίου και για κάθε ένα από αυτά υπολογίζεται η πιθανότητα να είναι το συγκεκριμένο σημείο ενδιαφέροντος αυτό στο οποίο πραγματοποιήθηκε η δραστηριότητα. Η πιθανότητα είναι αντιστρόφως ανάλογη προς την απόσταση του σημείου από το κεντροειδές του MBR του επεισοδίου και υπολογίζεται με βάση τον παρακάτω ψευδοκώδικα.

Αλγόριθμος pois_probability (MBR Episode MBR, D PoisTable)

// Υπολόγισε την συνολική απόσταση όλων των σημείων ενδιαφέροντος του πίνακα D, από το κεντροειδές του MBR, για τα όποια ισχύει ότι βρίσκονται εντός του ή εφάπτονται σε αυτό.
 $sum_dist = cacl_dist_sum (relate(D, CENTROID_{MBR}))$;

ΓΙΑ ΚΑΘΕ σημείο ενδιαφέροντος P_i του D το οποίο βρίσκεται εντός ή εφάπτεται στο δοθέν MBR
 ΕΠΑΝΑΛΑΒΕ

// Υπολόγισε το αποτέλεσμα της διαίρεσης της συνολικής απόστασης sum_dist , με την απόσταση από το κεντροειδές του MBR κάθε σημείου P_i , του πίνακα D, το οποίο βρίσκεται εντός του MBR ή εφάπτεται σε αυτό.
 $div_POI_i = (sum_dist / dist_POI_CENTROID (P_i))$;

ΓΙΑ ΚΑΘΕ-ΤΕΛΟΣ

// Υπολόγισε το άθροισμα των παραπάνω αποτελεσμάτων.
 $sum_div = sum (div_POI_i)$;

ΓΙΑ ΚΑΘΕ σημείο ενδιαφέροντος P_i του D το οποίο βρίσκεται εντός ή εφάπτεται στο δοθέν MBR
 ΕΠΑΝΑΛΑΒΕ

// Βρες την πιθανότητα να είναι αυτό το σημείο στο οποίο πραγματοποιήθηκε το επεισόδιο.
 $prob_POI_i = div_POI_i / sum_div$;

ΓΙΑ ΚΑΘΕ-ΤΕΛΟΣ

Επίσης στην προσθήκη αυτή δίνεται η δυνατότητα επιλογής απεικόνισης ή όχι των αποτελεσμάτων.

4.3.4 Δημιουργία μήτρας προέλευσης-προορισμού.

Οι μήτρες προέλευσης-προορισμού αποτελούν μέρος μικροσκοπικών μοντέλων κίνησης, μοντέλων πρόβλεψης κίνησης αλλά και αποθηκών δεδομένων τροχιών. Μπορούν να παραχθούν από δεδομένα κίνησης με την τμηματοποίηση του εξεταζόμενου χώρου σε διακριτές περιοχές ή σε επικαλυπτόμενες περιοχές ή σε κελιά $N \times N$ διαστάσεων. Με βάση αυτές, ακολουθεί ο υπολογισμός των τιμών, οι οποίες χρησιμοποιούνται για να απαντηθούν ερωτήματα του τύπου:

ποια (ή πόσα) κινούμενα αντικείμενα ξεκίνησαν από μια περιοχή A και σταμάτησαν σε μια περιοχή B;

ποια (ή πόσα) κινούμενα αντικείμενα επικαλύπτονται με μια περιοχή A και στην συνέχεια με μια περιοχή B κατά την διάρκεια της πορείας τους;

	A	B	C	D
A	1		1	
B			1	2
C				
D				

Εικόνα 7: Απεικόνιση μιας μήτρας προέλευσης-προορισμού

Στην περίπτωση μας ακολουθείται η προσέγγιση διαμοιρασμού του χώρου σε κελιά NxN και στη συνέχεια εφαρμόζονται δύο διαφορετικοί αλγόριθμοι για τον υπολογισμό των τιμών α) από τις σημασιολογικά επαυξημένες τροχιές και β) από την αντίστοιχη αποθήκη δεδομένων. Οι αλγόριθμοι που χρησιμοποιούνται δείχνονται παρακάτω.

α)

```

Αλγόριθμος construct_odmatrix (D MODSemanticTrajectoryTable)
// Δημιούργησε την δομή ενός πίνακα Προέλευσης-Προορισμού με βάση τα κελιά διαχωρισμού της
περιοχής
ODmatrix = populate_odmatrix_structure();
// Για κάθε semantic trajectory
ΓΙΑ ΚΑΘΕ Semantic Trajectory STi του D
ΕΠΑΝΑΛΑΒΕ
    // Βρες το MBR του πρώτου και τελευταίου επεισοδίου
    (MBRFrstEp, MBRLstEp) = Find_FirstLast_MBR(STi);
    // Βρες σε ποια κελιά της διαχωρισθείσας περιοχής περιέχονται τα MBRs αυτών των
επεισοδίων
    (CellO, CellD) = Contains(MBRFrstEp, MBRLstEp);
    // Αύξησε κατά ένα την τιμή στο αντίστοιχο πεδίο του πίνακα Προέλευσης-Προορισμού
    ODmatrix(CellO, CellD) := +1;
ΓΙΑ ΚΑΘΕ-ΤΕΛΟΣ

```

Εικόνα 8: Αλγόριθμος construct_odmatrix

β)

```

Αλγόριθμος construct_dw_odmatrix (D DWMovesFactTable)
// Δημιούργησε την δομή ενός πίνακα Προέλευσης-Προορισμού με βάση τα κελιά διαχωρισμού της
περιοχής
ODmatrix = populate_odmatrix_structure();
// Για κάθε επεισόδιο MOVE
ΓΙΑ ΚΑΘΕ επεισόδιο MOVE Mi του D
ΕΠΑΝΕΛΑΒΕ
    // Βρες το MBR των επεισοδίων STOP-ΑΠΟ και STOP-ΠΡΟΣ που περικλείουν το επεισόδιο
Mi
    (MBRFromEp, MBRToEp) = Find_FromTo_MBR(Mi);
    // Βρες σε ποια κελιά της διαχωρισθείσας περιοχής περιέχονται τα MBRs αυτών των
επεισοδίων

```

```

(Cello, CellD)=Contains(MBRFromEp, MBRToEp);
// Αύξησε κατά ένα την τιμή στο αντίστοιχο πεδίο του πίνακα Προέλευσης-Προορισμού
ODmatrix(Cello, CellD):= +1;
ΓΙΑ ΚΑΘΕ-ΤΕΛΟΣ

```

Εικόνα 9: Αλγόριθμος construct_dw_odmatrix

Η κύρια διαφορά των δύο παραπάνω αλγορίθμων είναι το σύνολο δεδομένων στο οποίο βασίζονται για να εξάγουν το αποτέλεσμα. Ενώ στην πρώτη περίπτωση ο υπολογισμός γίνεται στην προέλευση - προορισμό της κάθε σημασιολογικής τροχιάς, στη δεύτερη περίπτωση εφαρμόζεται στην προέλευση-προορισμό κάθε MOVE επεισοδίου. Για να υπάρχει λοιπόν η δυνατότητα σύγκρισης παρόμοιων συνόλων δεδομένων δημιουργήθηκε και μια λειτουργία μετασχηματισμού της κάθε τροχιάς από ακολουθία επεισοδίων STOPS και MOVES σε πολλές τροχιές με αρχή και τέλος STOP επεισόδια και ενδιάμεσα τα MOVE επεισόδια που μεσολαμβάνουν.

Για παράδειγμα μια τροχιά της μορφής $SEM_TRAJ = STOP:HOME \rightarrow MOVE:CAR \rightarrow STOP:WORK \rightarrow MOVE:CAR \rightarrow STOP:HOME$ μπορεί να μετασχηματιστεί στις αντίστοιχες δύο τροχιές $SEM_TRAJ1 = STOP:HOME \rightarrow MOVE:CAR \rightarrow STOP:WORK$ και $SEM_TRAJ2 = STOP:WORK \rightarrow MOVE:CAR \rightarrow STOP:HOME$. Κατά αυτό τον τρόπο τα παραγόμενα δεδομένα συμβαδίζουν με τα δεδομένα του πίνακα γεγονότων των MOVE επεισοδίων στην αποθήκη δεδομένων και έτσι μπορεί εύκολα να γίνει σύγκριση των δύο αλγορίθμων.

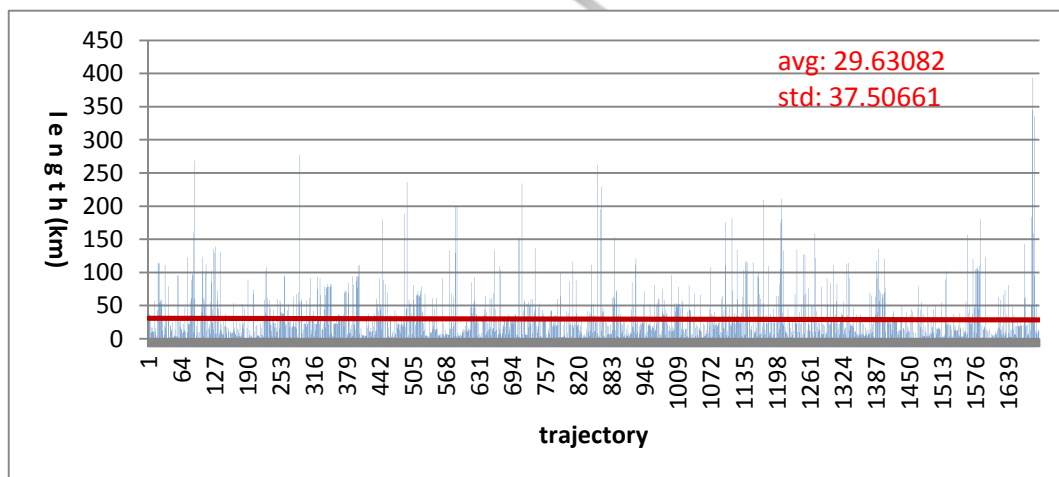
5 Παράδειγμα χρήσης

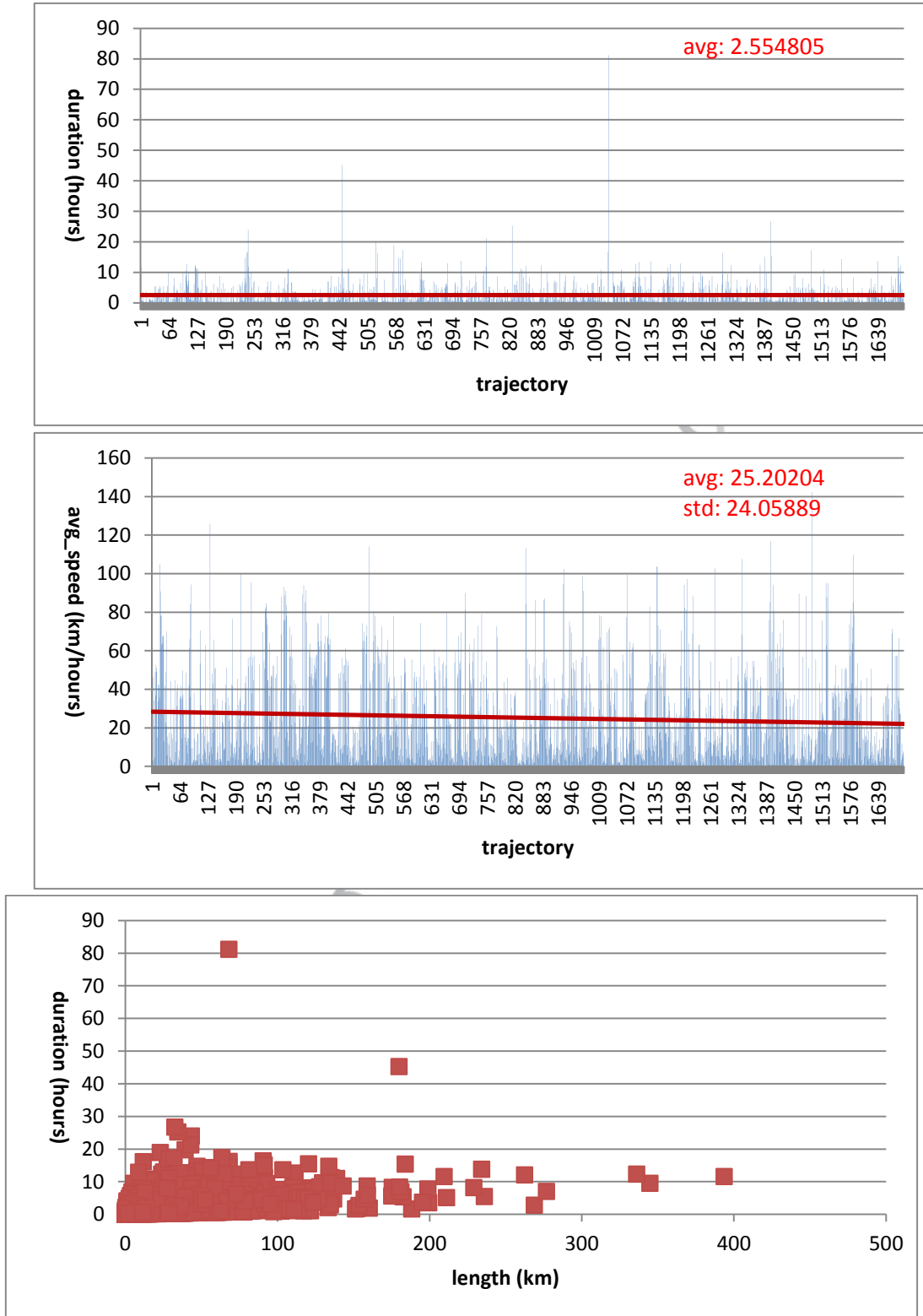
Στο κεφάλαιο αυτό θα παρουσιαστούν όσα περιγράφηκαν στο προηγούμενο κεφάλαιο πάνω σε ένα πραγματικό σύνολο δεδομένων σημασιολογικά επαυξημένων τροχιών πάντα μέσα από το περιβάλλον του HERMES MOD και με τη βοήθεια αρχείων KML και της εφαρμογής Google Earth, όπου απαιτείται οπτικοποίηση των αποτελεσμάτων.

5.1 Το δοθέν σύνολο δεδομένων

Το αρχικό σύνολο προέρχεται από την περιοχή του Βελγίου και περιλαμβάνει δύο πηγές δεδομένων, τα “GPS δεδομένα” τα οποία είναι οι gps εγγραφές ανά χρήστη, τα οποία εξήχθησαν από συσκευές με δυνατότητα εντοπισμού θέσης και τα “Ημερολόγια”, στα οποία κάθε χρήστης πρόσθεσε τη σημασιολογική πληροφορία της κάθε μετακίνησης από ένα σημείο Α σε ένα σημείο Β για συγκεκριμένο σκοπό. Και τα δύο σύνολα αφορούν μια περίοδο 730 ημερών από την 13-4-2006 έως την 14-4-2008.

Για την ανακατασκευή των τροχιών από τις έγκυρες δειγματοληπτιμένες εγγραφές χρησιμοποιήθηκε ένας απλός αλγόριθμος ο οποίος διέτρεξε για κάθε χρήστη τα ταξινομημένα δείγματα λαμβάνοντας υπόψη κάποιο χωρικό και χρονικό κατώφλι απόστασης. Με αυτό τον τρόπο πρόεκυψαν συνολικά 1698 τροχιές οι οποίες αντιστοιχούν σε 3542194 δείγματα με μέσο όρο 9,129 τροχιές ανά χρήστη και συνολικό μήκος 50313133,007449 το οποίο αντιστοιχεί σε 29630,82038 μέτρα ανά τροχιά. Επίσης ο μέσος όρος ταχύτητας υπολογίστηκε στα 7,000565 m/s ενώ το Ελάχιστο Περιβάλλον Ορθογώνιο όλων των τροχιών καθορίζεται από τις συντεταγμένες [(2.598423, 49.926288), (5.878673, 52.351493)]. Τα παραπάνω αποτελέσματα/στατιστικά δείχνονται στα γραφήματα που ακολουθούν.



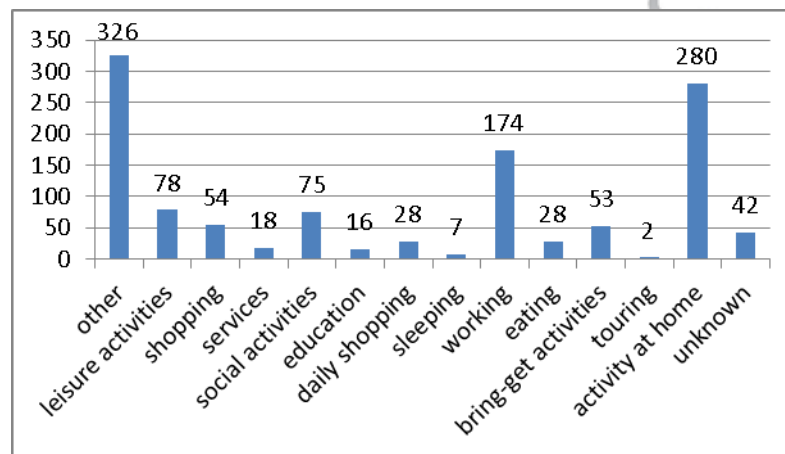


Εικόνα 10: Γραφήματα δεδομένων των ανακατασκευασμένων τροχιών από gps εγγραφές.

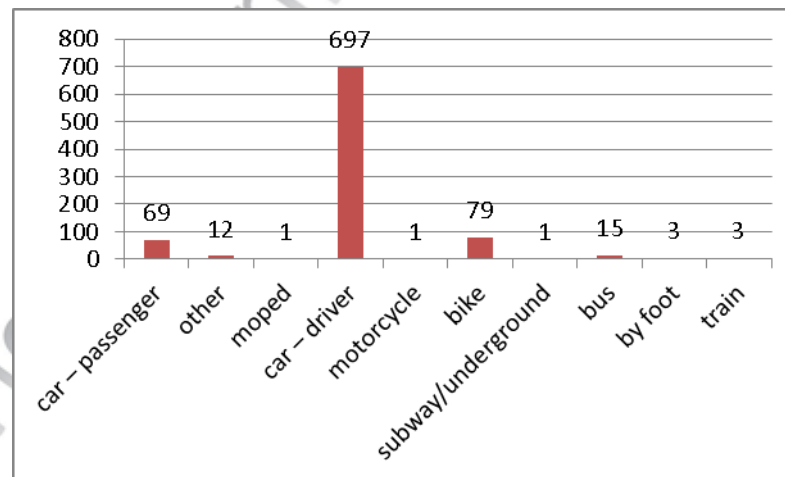
Κατά τη σύνθεση τώρα των σημασιολογικά επαυξημένων τροχιών έγινε χρήση των “Ημερολογίων” τα οποία όπως είπαμε και παραπάνω περιέχουν τις μετακινήσεις και την αντίστοιχη

σημασιολογική πληροφορία αυτών. Τέτοιου είδους πληροφορία είναι τα σημείο εκκίνησης και προορισμού, η δραστηριότητα στον προορισμό καθώς και το χρησιμοποιούμενο μεταφορικό μέσο. Υπάρχουν συνολικά 300 μετακινήσεις ανά μέρα και ανά χρήστη και κάθε μια από αυτές μπορεί να συνδεθεί μέσω ενός αναγνωριστικού αριθμού με τα gps δεδομένα που παρήχθησαν κατά την διάρκεια της. Στην προσέγγιση που χρησιμοποιήθηκε μια τροχιά προκύπτει κάθε φορά που συναντάμε εγγραφή για νέο χρήστη ή για νέα μέρα. Κατ' αυτό τον τρόπο προκύπτουν 300 σημασιολογικές τροχιές.

Επειδή μια μετακίνηση θεωρείται MOVE επεισόδιο εισήχθησαν STOP επεισόδια πριν και μετά από αυτό. Δημιουργήθηκαν λοιπόν 881 MOVE και 1181 STOP επεισόδια με μέσο όρο 6.87333 επεισόδια για κάθε τροχιά. Τα MOVE και STOP επεισόδια μπορούμε να τα ταξινομήσουμε βασιζόμενοι στην σημασιολογική πληροφορία που διαθέτουν, έτσι διακρίνουμε 10 κατηγορίες MOVE επεισοδίων ως προς τα χρησιμοποιηθέντα μεταφορικά μέσα και 14 κατηγορίες STOP επεισοδίων ως προς τις δραστηριότητες στον προορισμό των οποίων αντίστοιχα γραφήματα φαίνονται παρακάτω.

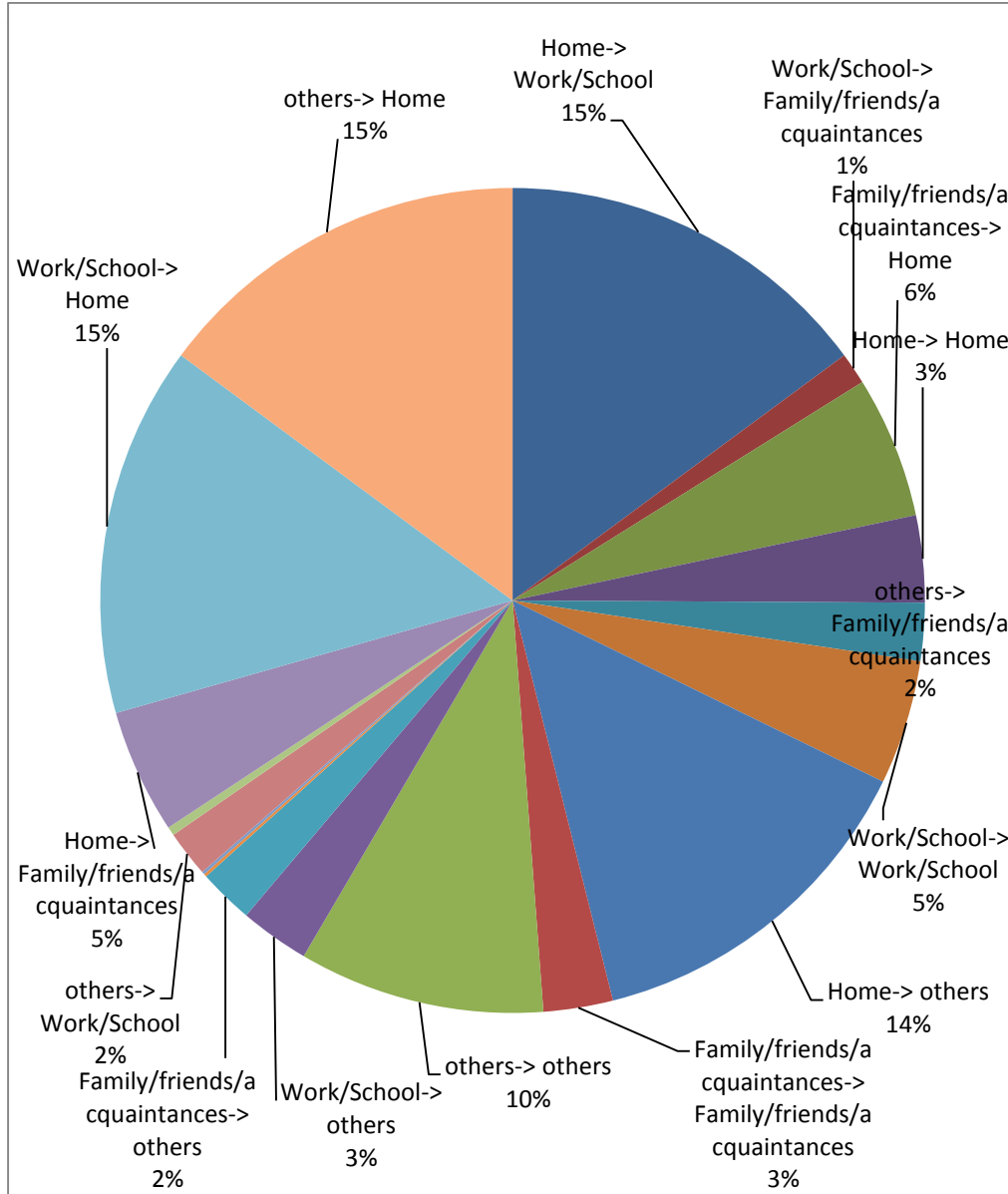


Εικόνα 11: Κατανομή των STOP επεισοδίων ως προς τον σκοπό στον προορισμό.



Εικόνα 12: Κατανομή των MOVE επεισοδίων ως προς το χρησιμοποιηθέν μεταφορικό μέσο.

Επίσης στο παρακάτω γράφημα φαίνεται η κατανομή των MOVE επεισοδίων ως προς το είδος εκκίνησης - προορισμού.



Στη συνέχεια για την περαιτέρω επεξεργασία του συνόλου δεδομένων ο αλγόριθμος T-OPTICS ο οποίος περιγράφηκε στην παράγραφο 0 τροφοδοτήθηκε με όλες τις τροχιές οι οποίες στην ουσία αντιστοιχούν στα MOVE επεισόδια. Παρήχθησαν έτσι για κάθε τροχιά ενδιάμεσα σημεία ακινησίας, δηλαδή STOP επεισόδια, με αποτέλεσμα ο συνολικός αριθμός των επεισοδίων να είναι πλέον 4634 με 2530 από αυτά να είναι STOP και 2104 MOVE. Παράλληλα από την αξιοποίηση των gps δεδομένων παρήχθησαν οι σχετικές υπο-τροχιές του κάθε επεισοδίου.

Τέλος στο δοθέν σύνολο τα σημεία ενδιαφέροντος (Points Of Interest) που εμπεριέχονται βρέθηκε ότι αφορούν 289 δήμους και μπορούν να κατηγοριοποιηθούν σε 4 ομάδες Home, Work/School, Family/friends/acquaintances, others.

5.2 Εξερεύνηση δυνατοτήτων HERMES MOD

Για την καλύτερη κατανόηση των ερωτημάτων κρίνεται σκόπιμο να γίνει μια ποιο πλήρης περιγραφή των συναρτήσεων και των τύπων του HERMES που θα χρησιμοποιηθούν στην επόμενη ενότητα. Όπως έχουμε αναφέρει το υπάρχον σχήμα περιλαμβάνει δύο πίνακες αντικειμένων, οι οποίοι περιέχουν όλοι

την παραγόμενη πληροφορία. Αυτοί είναι οι 'BELG_SEM_TRAJS' and 'BELG_SUB_MPOINTS' για τις σημασιολογικά επαυξημένες τροχιές και τα MOVING POINTS (που είναι το τύπος για την τροχιά στο HERMES) αντίστοιχα.

Σε κάθε αντικείμενο BELG_SUB_MPOINT εμπεριέχονται μέθοδοι που κληρονομούνται από το αντικείμενο MOVING_POINT καθώς και η μέθοδος *getsemmbb()* η οποία επιστρέφει ένα SEM_MBB αντικείμενο το θεωρητικό του υπόβαθρο του οποίου έχει περιγράφει στη παράγραφο 4.1.1. Με την σειρά του το SEM_MBB εμπεριέχει τις εξής μεθόδους:

area(srid): επιστρέφει την χωρική περιοχή του MBB με *srid* το αναγνωριστικό του χωρικού συστήματος αναφοράς,

duration(): επιστρέφει την χρονική περιοχή του MBB,

getrectangle(srid): επιστρέφει το ελάχιστο περιβάλλον ορθογώνιο (MBR) με την μορφή αντικείμενου MDSYS.GEOMETRY που είναι ο γεωμετρικός τύπος του Oracle Spatial και με *srid* το αναγνωριστικό του χωρικού συστήματος αναφοράς.

Το SEM_EPISODE περιέχει τις μεθόδους *duration()*: επιστρέφει, βασιζόμενο στο MBB, την χρονική διάρκεια του επεισοδίου.

Το SEM_TRAJECTORY περιλαμβάνει τις:

num_of_stops()/num_of_moves: επιστρέφει τον αριθμό των εμπεριεχομένων STOP/MOVE επεισοδίων,

getMBB(): επιστρέφει το MBB της σημασιολογικής τροχιάς,

sem_stops()/sem_moves(): επιστρέφει όλα τα STOP/MOVE επεισόδια της τροχιάς με την μορφή πίνακα,

at(spatial,temporal,episode_tag,activity_tag): επιστρέφει μια σημασιολογική τροχιά περιορισμένη στις δοθείσες παραμέτρους,

confined_in(geometry,period,tag): επιστρέφει μια σημασιολογική τροχιά η οποία τέμνεται με το γεωμετρικό περιορισμό *geometry*, με το χρονικό περιορισμό *period* (τύπος *tau_tll.d_period_sec* - (N.Pelekis, 2002)) και τις δοθείσες ετικέτες *tag* οι οποίες μπορούν να είναι κείμενο έως 1000 χαρακτήρες της μορφής 1 "*tag1+tag2+...+tagn*",

episodes_with(tag): επιστρέφει το σύνολο των επεισοδίων μιας τροχιάς τα οποία έχουν ετικέτες παρόμοιες με τις ετικέτες εισόδου οι οποίες μπορούν να είναι κείμενο έως 1000 χαρακτήρες της μορφής 1 "*tag1+tag2+...+tagn*".

Πρόσθετα υπάρχουν οι ακόλουθες διαδικασίες για την εκτέλεση ερωτημάτων εύρους βασιζόμενα σε ένα τύπο ευρετηρίου που έχει αναπτυχθεί ειδικά για σημασιολογικά επαυξημένες τροχιές το οποίο ονομάζεται STB-TREE, και είναι οι εξής:

- *std_range_episodes(episodeType,geometry)*: επιστρέφει *episodeType* επεισόδια τα οποία διασταυρώνονται με το δοθέντα γεωγραφικό περιορισμό (*geometry*)
- *std_range_episodes(episodeType,temporalPeriod)*: επιστρέφει *episodeType* επεισόδια τα οποία διασταυρώνονται με το δοθέντα χρονικό περιορισμό (*temporalPeriod*).
- *std_range_episodes(fromGeometry,toGeometry, temporalPeriod)*: επιστρέφει MOVE επεισόδια τα οποία ξεκίνησαν (τερμάτισαν) από το (στο) *fromGeometry(toGeometry)* κατά την χρονική περίοδο *temporalPeriod*.
- *std_range_episodes_mbr(fromStop,toStop, viaMove)*: επιστρέφει όλα τα MOVE επεισόδια τα οποία ξεκίνησαν (τερμάτισαν) από το (στο) *fromStop (toStop)* περνώντας από το *viaMove* επεισόδιο το οποίο μπορεί να είναι και null.

5.3 Ερωτήματα στο HERMES MOD

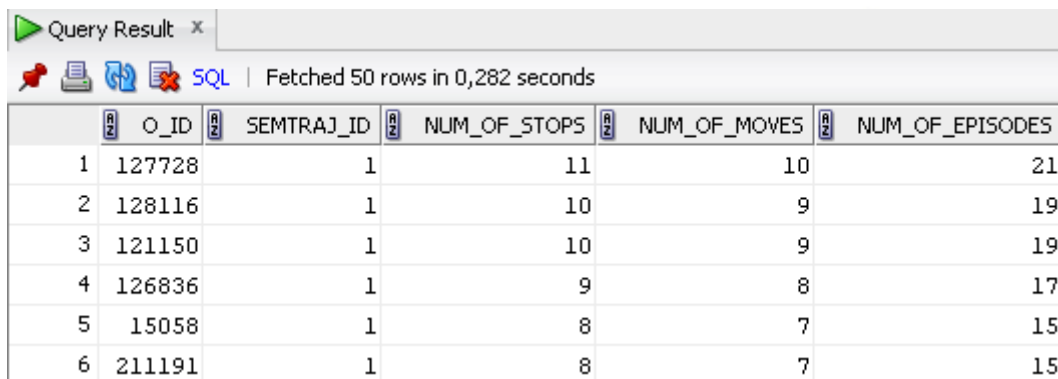
5.3.1 Εφαρμογή τελεστών για τον υπολογισμό του αριθμού των επεισοδίων

Σκοπός της παρακάτω συνάρτησης είναι να υπολογίσει τον αριθμό Stops και Moves για όλες τις σημασιολογικές τροχιές του πίνακα.

```

SELECT o_id,
       semtraj_id,
       VALUE (t).num_of_stops () num_of_stops,
       VALUE (t).num_of_moves () num_of_moves,
       (VALUE (t).num_of_stops () + VALUE (t).num_of_moves () ) AS
num_of_episodes
FROM belg_sem_trajs t
ORDER BY 4 DESC;

```



	O_ID	SEMTRAJ_ID	NUM_OF_STOPS	NUM_OF_MOVES	NUM_OF_EPISODES
1	127728	1	11	10	21
2	128116	1	10	9	19
3	121150	1	10	9	19
4	126836	1	9	8	17
5	15058	1	8	7	15
6	211191	1	8	7	15

Εικόνα 13: Αποτελέσματα ερωτήματος υπολογισμού αριθμού επεισοδίων

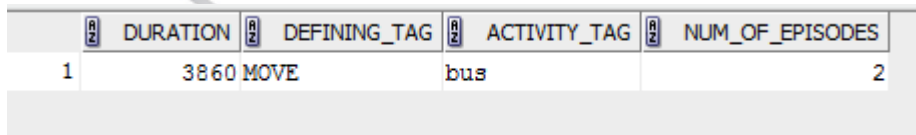
5.3.2 Χρήση τελεστή για την εύρεση επεισοδίων

Σε αυτό το ερώτημα γίνεται χρήση του τελεστή *episodes_with* και στην συνέχεια υπολογίζεται το άθροισμα της διάρκειας των επεισοδίων MOVE ανά τρόπο μεταφοράς, για δοθείσα σημασιολογική τροχιά. Στην περίπτωση αυτή δίδεται σαν όρισμα η ετικέτα 'MOVE'.

```

SELECT SUM (value(s).duration(). m_Value ) duration,
       defining_tag, activity_tag, COUNT (defining_tag)
num_of_episodes
FROM TABLE (SELECT t.episodes_with ('MOVE')
             FROM belg_sem_trajs t
             WHERE t.o_id = 5238 AND t.semtraj_id = 1) s
GROUP BY defining_tag, activity_tag;

```



	DURATION	DEFINING_TAG	ACTIVITY_TAG	NUM_OF_EPISODES
1	3860	MOVE	bus	2

Εικόνα 14: Αποτελέσματα ερωτήματος

5.3.3 Σύνθετη χρήση τελεστών πάνω στην σημασιολογική πληροφορία

Γίνεται εντοπισμός επεισοδίων MOVE, των οποίων η διάρκεια είναι μικρότερη από τη μέση διάρκεια των επεισοδίων MOVE με δραστηριότητα 'working'. Η λειτουργία αυτή δείχνει πώς μπορεί κανείς να φιλτράρει τα επεισόδια με πολλαπλές ετικέτες.

```
SELECT VALUE (ext_s).DURATION () .m_value DURATION,
        defining_tag,
        activity_tag
FROM TABLE
    (SELECT t.isodes_with ('MOVE')
    FROM belg_sem_traj_s t
    WHERE t.o_id      = 5238
    AND t.semtraj_id = 2
    ) ext_s
WHERE (VALUE (ext_s).DURATION () .m_value) <
    (SELECT AVG (VALUE (s).DURATION () .m_value)
    FROM TABLE
        (SELECT t.isodes_with ('working')
        FROM belg_sem_traj_s t
        WHERE t.o_id      = 5238
        AND t.semtraj_id  = 2
        AND t.isodes_with ('MOVE') IS NOT NULL
        ) s
    );
```

ID	DURATION	DEFINING_TAG	ACTIVITY_TAG
1	1022	MOVE	subway/underground
2	1016	MOVE	by foot
3	2620	MOVE	bus
4	3211	MOVE	car - passenger

Εικόνα 15: Αποτελέσματα ερωτήματος με χρήση φίλτρου πολλαπλών ετικετών

5.3.4 Περιορισμός στη χωρική διάσταση, φιλτράρισμα με βάση τη σημασιολογική πληροφορία κειμένου και οπτικοποίηση των αποτελεσμάτων.

Περιορίζουμε μια δοθείσα σημασιολογική τροχιά μέσα σε μια χρονική περίοδο μέσω της *confined_in* και στη συνέχεια επιστρέφουμε μόνο τα επεισόδια STOP κατά τα οποία ο χρήστης εργαζόταν μέσω της Διαχείριση σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων

episodes_with. Επίσης, παρουσιάζεται και χρήση της επέκτασης (*visualizer.semtrajectory2kml*) για την οπτικοποίηση των αποτελεσμάτων. Ο ορισμός της χρονικής περιόδου χρήσης γίνεται με τον τύπο `tau_tll.d_period_sec`.

```
SELECT *
FROM TABLE (
    SELECT b.confined_in (NULL,
        tau_tll.d_period_sec (
            tau_tll.d_timepoint_sec (2007, 10, 22, 08, 00, 00 ),
            tau_tll.d_timepoint_sec (2007, 10, 22, 23, 59, 00 )
        ), 'STOP' ).episodes_with ('working')
    FROM belg_sem_traj_s b
    WHERE b.o_id = 216828
    AND b.semtraj_id = 2
);
```

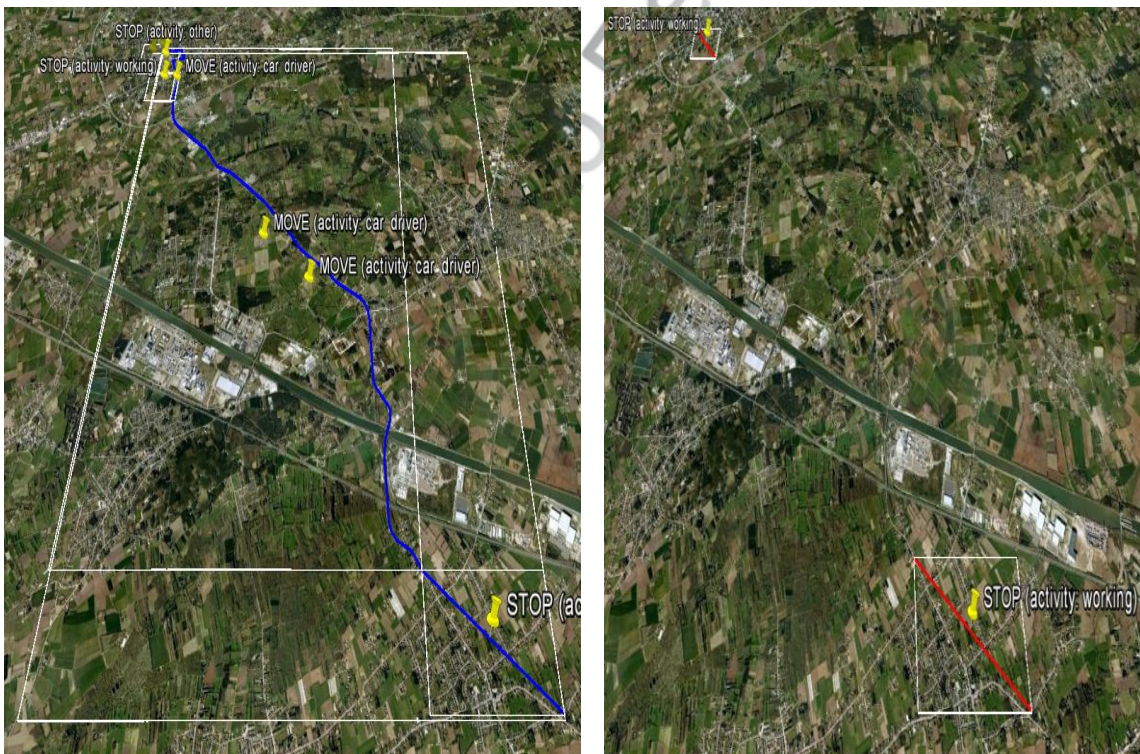
Οπτικοποίηση

```
DECLARE
    vsrid          INTEGER := 4326;
    i              INTEGER := 0;
    void           INTEGER := 216828;
    vsemtrajid    INTEGER := 2;
    sb_mps mp_array := mp_array ();
    semtraj sem_trajectory;
BEGIN
    FOR rc IN
        (SELECT Deref (tlink).sub_mpoint sub_mpoint
        FROM TABLE
            (SELECT b.confined_in (NULL, tau_tll.d_period_sec
            (tau_tll.d_timepoint_sec (2007, 10, 22, 08, 00, 00 ),
            tau_tll.d_timepoint_sec (2007, 10, 22, 23, 59, 00 ) ), 'STOP'
            ).episodes_with ('working')
            FROM belg_sem_traj_s b
            WHERE b.o_id = void
            AND b.semtraj_id = vsemtrajid
            ) s
```

```

)
LOOP
  sb_mps := mp_array (rc.sub_mpoint);
  visualizer.movingpointtable2kml (sb_mps, vsrid, 'u' || void || '_'
|| i || '_MOVPOINT.kml' );
  i := i + 1;
END LOOP;
SELECT VALUE (s)
INTO semtraj
FROM belg_sem_traj s
WHERE o_id = void
AND semtraj_id = vsemtrajid;
visualizer.semtrajectory2kml (semtraj, 'TRUE', 'TRUE', 'TRUE');
END;

```



Εικόνα 16: Οπτικοποίηση αποτελεσμάτων. Στις εικόνες φαίνονται, με μπλε χρώμα, το semantic trajectory στο σύνολο του και με κόκκινο χρώμα τα επεισόδια που επιστρέφει το ερώτημα.

Αυτό το ερώτημα είναι μια παραλλαγή του προηγούμενου, στο οποίο περιορίζεται μια δοθείσα σημασιολογική τροχιά μέσα σε μια χρονική περίοδο και στη συνέχεια επιστρέφει STOP είτε WORKING επεισόδια (ετικέτες στην *confined_in*).

```

SELECT *
  FROM TABLE (
SELECT b.confined_in (NULL,
tau_tll.d_period_sec (
tau_tll.d_timepoint_sec (2007, 10, 22, 08, 00, 00 ),
tau_tll.d_timepoint_sec (2007, 10, 22, 23, 59, 00 )
), 'STOP+working' ).episodes
FROM belg_sem_traj_s b
WHERE b.o_id = 216828
AND b.semtraj_id = 2
);

```

Οπτικοποίηση

```

DECLARE
vsrid      INTEGER := 4326;
i          INTEGER := 0;
void       INTEGER := 216828;
vsemtrajid INTEGER := 2;
sb_mps mp_array := mp_array ();
semtraj sem_trajectory;
BEGIN
FOR rc IN
(SELECT Deref (tlink).sub_mpoint sub_mpoint
FROM TABLE
(SELECT b.confined_in (NULL,
tau_tll.d_period_sec (
tau_tll.d_timepoint_sec (2007, 10, 22, 08, 00, 00 ),
tau_tll.d_timepoint_sec (2007, 10, 22, 23, 59, 00 )
), 'STOP+working' ).episodes
FROM belg_sem_traj_s b
WHERE b.o_id = void
AND b.semtraj_id = vsemtrajid
) s
)
LOOP

```



```

sb_mps := mp_array (rc.sub_mpoint);

visualizer.movingpointtable2kml (sb_mps, vsrid, 'u' || void || '_'
|| i || '_MOVPOINT.kml' );

i := i + 1;

END LOOP;

SELECT VALUE (s)

INTO semtraj

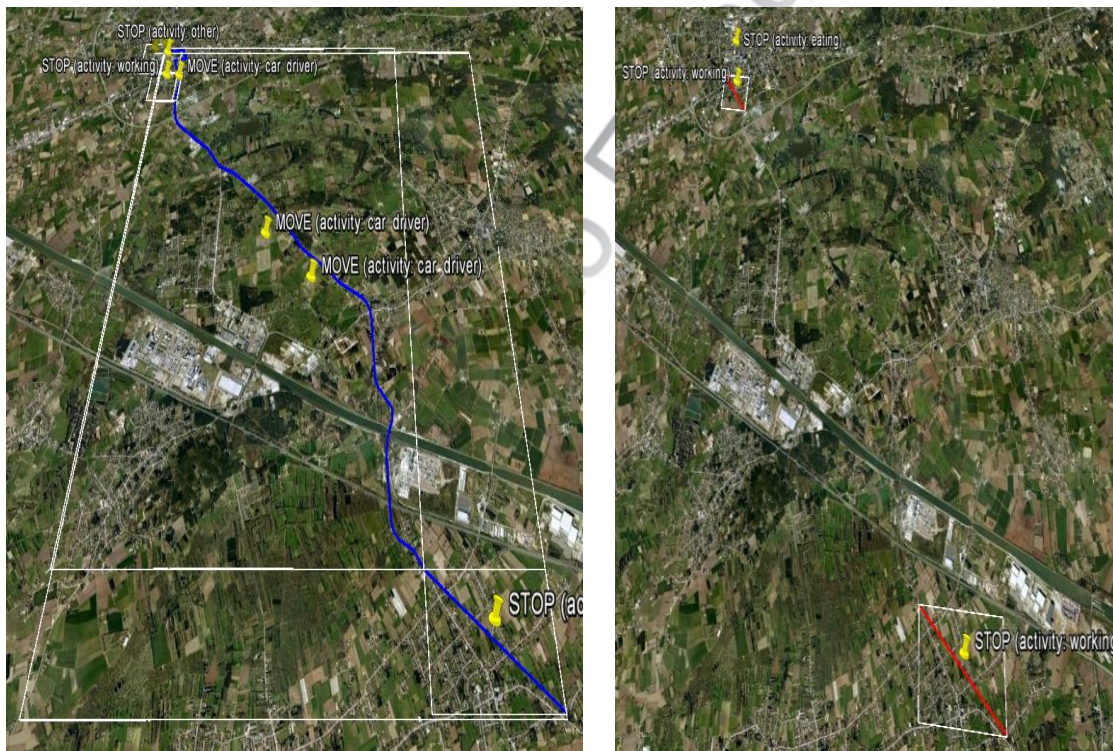
FROM belg_sem_traj_s

WHERE o_id      = void

AND semtraj_id = vsemtrajid;

visualizer.semtrajectory2kml (semtraj, 'TRUE', 'TRUE', 'TRUE');

END;
```



Εικόνα 17: Οπτικοποίηση αποτελεσμάτων. Στις εικόνες φαίνονται, με μπλε χρώμα, το semantic trajectory στο σύνολο του. Με κόκκινο χρώμα, τα επεισόδια που επιστρέφει το ερώτημα.

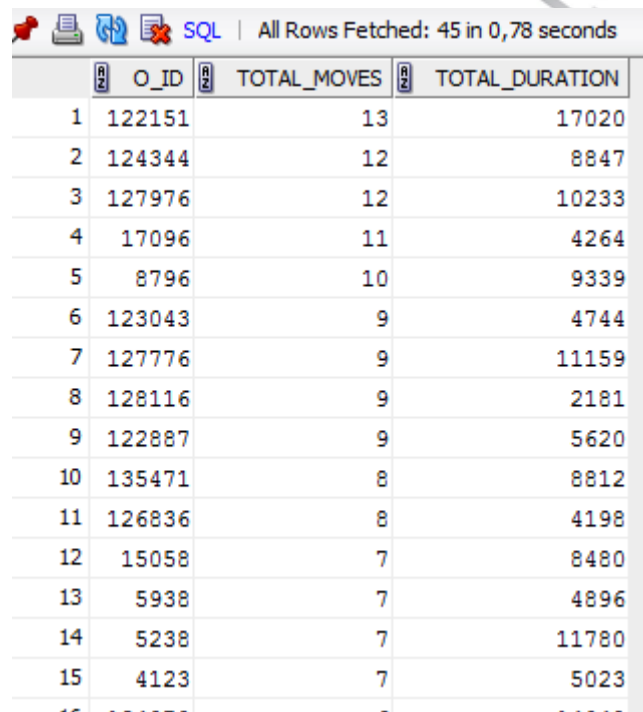
5.3.5 Ερώτημα χρονικού εύρους και φίλτρου στην σημασιολογική πληροφορία με χρήση ευρετηρίου.

Στην περίπτωση αυτή επιστρέφεται ο αριθμός μόνο των επεισοδίων MOVE, και υπολογίζεται γι αυτά το άθροισμα της διάρκειάς τους. Εδώ γίνεται χρήση της `std_range_episodes(episodeType,temporalPeriod)`.

```

SELECT Deref (tlink).o_id o_id,
       COUNT (tlink) total_moves,
       SUM (value(s).duration(). m_value) total_duration
FROM TABLE
      (SELECT      std.stb_range_episodes      ('MOVE',      tau_tll.d_period_sec
(tau_tll.d_timepoint_sec      (2006,      01,      01,      00,      01,      00      ),
tau_tll.d_timepoint_sec (2006, 12, 31, 23, 59, 00 ) ), 'SEM_INDX' )
      FROM DUAL
      ) s
GROUP BY Deref (tlink).o_id
ORDER BY 2 DESC;

```



	O_ID	TOTAL_MOVES	TOTAL_DURATION
1	122151	13	17020
2	124344	12	8847
3	127976	12	10233
4	17096	11	4264
5	8796	10	9339
6	123043	9	4744
7	127776	9	11159
8	128116	9	2181
9	122887	9	5620
10	135471	8	8812
11	126836	8	4198
12	15058	7	8480
13	5938	7	4896
14	5238	7	11780
15	4123	7	5023

Εικόνα 18: Αποτελέσματα ερωτήματος χρονικού εύρους

5.3.6 Ερώτημα χωρο-χρονικού εύρους και φίλτρου στη σημασιολογική πληροφορία με χρήση ευρετηρίου.

Με αυτό το ερώτημα χωροχρονικού εύρους γίνεται υπολογισμός της διάρκειας των «Work/School» και «working» STOP επεισοδίων που λαμβάνουν χώρα σε μια περιοχή για ένα χρονικό διάστημα. Εδώ γίνεται χρήση της *std_range_episodes(episodeType,geometry)*.

```

SELECT      SUM (VALUE (s).DURATION () .m_value) work_duration,
           Deref (tlink).o_id o_id
FROM TABLE

```

```

      (SELECT std.stb_range_episodes

```

Διαχείριση σημασιολογικά επηρεασμένων τροχιών κινούμενων αντικειμένων

```

('STOP',
MDSYS.SDO_GEOMETRY
(2003,
4236,
NULL,
MDSYS.sdo_elem_info_array
(1, 1003, 3),
MDSYS.sdo_ordinate_array
(4.985388, 51.152885, 5.994323, 52.158382)
),
tau_t11.d_period_sec (tau_t11.d_timepoint_sec
(2007, 10, 22, 08, 00, 00),
tau_t11.d_timepoint_sec
(2008, 10, 22, 23, 00, 00)
),
'SEM_INDX'
)
FROM DUAL
) s
WHERE episode_tag = 'Work/School' AND activity_tag = 'working'
GROUP BY Deref (tlink).o_id;

```

RZ	WORK_DURATION	RZ	O_ID
1	4643	218411	
2	36024	216828	
3	16011	130049	

Εικόνα 19: Αποτελέσματα ερωτήματος χωροχρονικού εύρους

5.3.7 Ερώτημα χωρο - χρονικού εύρους διασταύρωσης και φίλτρου στη σημασιολογική πληροφορία με χρήση ευρετηρίου καθώς και εφαρμογή χρονικού περιορισμού.

Στην περίπτωση αυτή περιορίζονται οι υπο-τροχιές των επεισοδίων STOP (που υπάρχουν μέσα σε ένα χωροχρονικό πλαίσιο), μέσα σε μια χρονική περίοδο. Εδώ γίνεται χρήση της *at* (*spatial,temporal,episode_tag,activity_tag*), ενώ μέσω του *tlink* λαμβάνονται οι υπο-τροχιές.

```

SELECT Deref (tlink).sub_mpoint sub_mpoint,
Deref (tlink).sub_mpoint.at_period (tau_t11.d_period_sec
(tau_t11.d_timepoint_sec (2007, 10, 22, 8, 00, 00) ),
tau_t11.d_timepoint_sec (2007, 10, 22, 10, 00, 00) ) )
restricted_sub_mpoint,

```

```

    Deref (tlink).o_id o_id
FROM TABLE
    (SELECT std.stb_range_episodes ('STOP', MDSYS.SDO_GEOMETRY (2003,
4326, NULL, MDSYS.sdo_elem_info_array (1, 1003, 3 ),
MDSYS.sdo_ordinate_array (4.981388, 51.152885, 4.994323, 51.162082 )
), tau_tll.d_period_sec (tau_tll.d_timepoint_sec (2007, 10, 22, 5, 00,
00 ), tau_tll.d_timepoint_sec (2007, 10, 28, 5, 00, 00 ) ), 'SEM_INDX'
)
FROM DUAL) s ;

```

Οπτικοποίηση

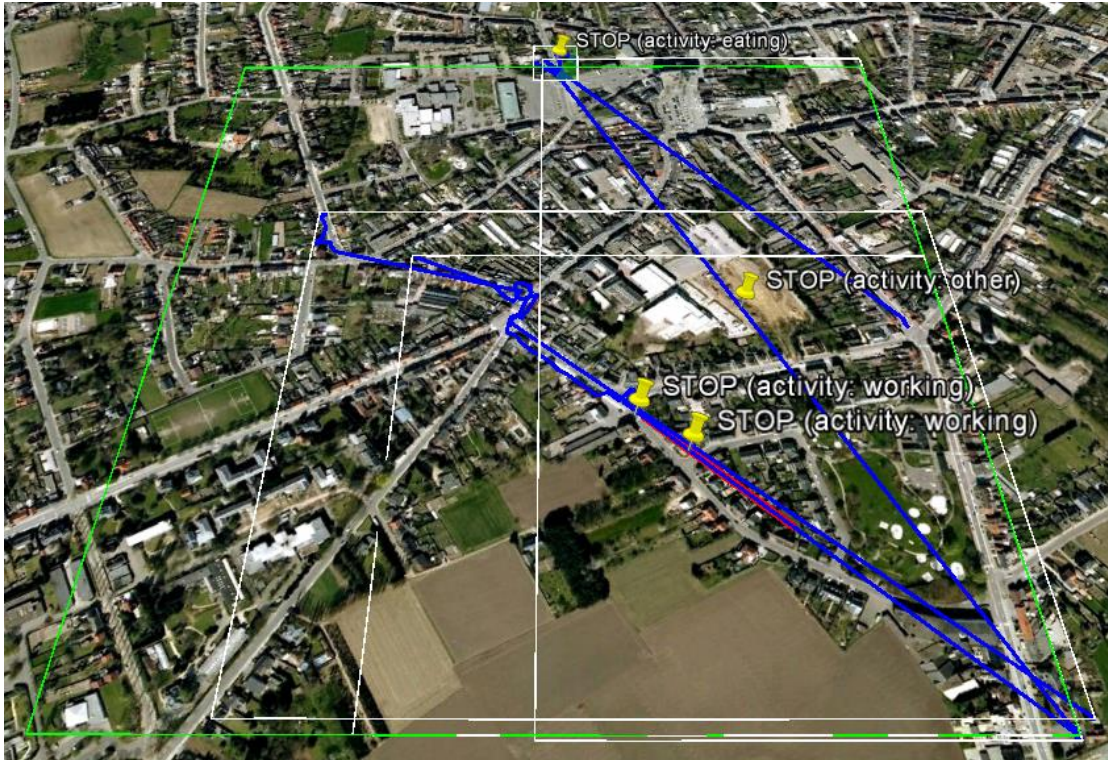
```

DECLARE
    vsrid INTEGER := 4326;
    i INTEGER := 0;
    sb_mps mp_array := mp_array ();
    rsb_mps mp_array := mp_array ();
    vgeom MDSYS.SDO_GEOMETRY := MDSYS.SDO_GEOMETRY (2003, vsrid, NULL,
MDSYS.sdo_elem_info_array (1, 1003, 3), MDSYS.sdo_ordinate_array
(4.981388, 51.152885, 4.994323, 51.162082 ) );
BEGIN
    FOR rc IN
        (SELECT Deref (tlink).sub_mpoint sub_mpoint,
            Deref (tlink).subtraj_id subtraj_id,
            Deref (tlink).traj_id traj_id,
            Deref (tlink).sub_mpoint.at_period (tau_tll.d_period_sec
(tau_tll.d_timepoint_sec (2007, 10, 22, 8, 00, 00 ) ),
tau_tll.d_timepoint_sec (2007, 10, 22, 10, 00, 00 ) ) )
        restricted_sub_mpoint,
            Deref (tlink).o_id o_id,
            MDSYS.sdo_geom.sdo_centroid (VALUE (s).mbb.getrectangle (vsrid),
0.01 ) mbr_centroid,
            VALUE (s).mbb.getrectangle (vsrid) mbb,
            activity_tag,
            episode_tag,
            defining_tag
        FROM TABLE
            (SELECT std.stb_range_episodes ('STOP', vgeom,
tau_tll.d_period_sec (tau_tll.d_timepoint_sec (2007, 10, 22, 5, 00, 00
), tau_tll.d_timepoint_sec (2007, 10, 28, 5, 00, 00 ) ), 'SEM_INDX' )

```



```
FROM DUAL
) s
)
LOOP
  sb_mps := mp_array (rc.sub_mpoint);
  IF rc.restricted_sub_mpoint IS NOT NULL THEN
    rsb_mps := mp_array
(rc.restricted_sub_mpoint);
    visualizer.movingpointtable2kml (rsb_mps, vsrid, 'u' || rc.o_id ||
'_ ' || i || '_RMOVPOINT.kml' );
  END IF;
  visualizer.movingpointtable2kml (sb_mps, vsrid, 'u' || rc.o_id ||
'_ ' || i || '_MOVPOINT.kml' );
  visualizer.placemark2kml (rc.mbr_centroid, vsrid, 'u' || rc.o_id
|| 'traj' || rc.traj_id || 'subtraj' || rc.subtraj_id ||
'_CENTROID.kml', rc.defining_tag || ' (activity: ' || rc.activity_tag
|| ')', rc.subtraj_id || '-' || rc.episode_tag );
  visualizer.polygon2kml (rc.mbb, vsrid, 'u' || rc.o_id || 'traj' ||
rc.traj_id || 'subtraj' || rc.subtraj_id || '_RECTANGLE.kml' );
  i := i + 1;
END LOOP;
visualizer.polygon2kml (vgeom, vsrid, 'RECTANGLE.kml');
END;
```



Εικόνα 20: Στην εικόνα φαίνονται, με μπλε χρώμα, τα επεισόδια με ετικέτα 'STOP' της δοθείσας χρονικής περιόδου και γεωμετρικού περιορισμού (περίγραμμα με πράσινο χρώμα). Με κόκκινο χρώμα, το τμήμα του επεισοδίου που βρίσκεται στην χρονική περίοδο που δηλώνετε κατά την χρήση του τελεστή at_period.

5.3.8 Εφαρμογή τελεστών για την εύρεση επεισοδίων τα οποία επικαλύπτονται με χρονική περίοδο.

Γίνεται εύρεση της μέσης διάρκειας εκείνων των επεισοδίων STOP που επικαλύπτονται με δοθείσα χρονική περίοδο.

```
SELECT SUM (VALUE (s).DURATION () .m_value) / COUNT (DEREF
(tlink).traj_id) avg_mbb_stop_duration
FROM TABLE
(SELECT std.stb_range_episodes ('STOP', MDSYS.SDO_GEOMETRY (2003,
4326, NULL, MDSYS.sdo_elem_info_array (1, 1003, 3 ),
MDSYS.sdo_ordinate_array (4.985388, 51.152885, 4.994323, 51.158382 )
), tau_tll.d_period_sec (tau_tll.d_timepoint_sec (2007, 10, 22, 08,
00, 00 ), tau_tll.d_timepoint_sec (2007, 10, 23, 08, 00, 00 ) ),
'SEM_INDX' )
FROM DUAL
) s
```

	AVG_MBB_STOP_DURATION
1	95863,5

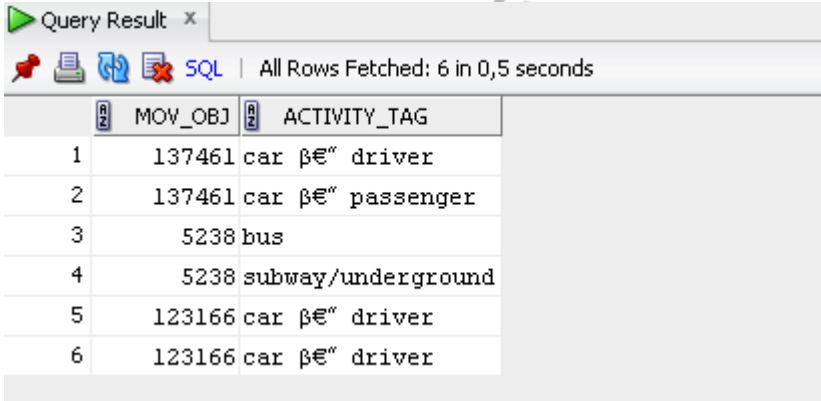
Εικόνα 21: Αποτελέσματα ερωτήματος

Διαχείριση σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων

5.3.9 Ερωτήματα εύρους σε ευρετήριο για την εύρεση μοτίβων της μορφής “από-προς” .

α) Ανακτώνται κινούμενα αντικείμενα (στην πραγματικότητα τα MOVE επεισόδια τους) που ξεκινούν από «Home» και καταλήγουν σε «Work/School» μεταξύ μιας δεδομένης χρονικής περιόδου. Χρήση της *std_range_episodes_mbr(fromStop,toStop, viaMove)*.

```
select      deref(tlink).o_id      mov_obj,      activity_tag      from
table(std.stb_range_episodes (
sem_episode ('STOP',      'Home',null,null,null),
sem_episode ('STOP',      'Work/School',null,null,null),
sem_episode ('MOVE',null,null,
sem_mbb(sem_st_point(5.503502,50.953662,
tau_tll.d_timepoint_sec(2006,5,1,5,00,00)),sem_st_point(5.727093,51.02
6928,tau_tll.d_timepoint_sec(2007,6,1,5,00,00))),
null),
'SEM_INDX')) t;
```



	MOV_OBJ	ACTIVITY_TAG
1	137461 car β€	driver
2	137461 car β€	passenger
3	5238 bus	
4	5238 subway/underground	
5	123166 car β€	driver
6	123166 car β€	driver

Εικόνα 22: Αποτελέσματα ερωτήματος

β) Πόσες φορές και με τι μεταφορικό μέσο ξεκινούν από «Home» που είναι μέσα σε μια δοθείσα περιοχή για μια δεδομένη χρονική περίοδο και η επόμενη στάση τους είναι «working».

```
SELECT COUNT(deref(tlink).o_id) sum_mov_obj,
activity_tag
FROM      TABLE(std.stb_range_episodes_mbr(      sem_episode ('STOP',
NULL,NULL,HERMES.SEM_MBB(HERMES.SEM_ST_POINT(4.938277,50.964135,TAU_TL
L.D_TIMEPOINT_SEC(2007,1,1,7,00,00)),HERMES.SEM_ST_POINT(5.503228,51.2
04038,TAU_TLL.D_TIMEPOINT_SEC(2008,1,1,7,00,00))),NULL),
sem_episode ('STOP', NULL,'working',NULL,NULL), NULL, 'SEM_INDX')) t
GROUP BY activity_tag;
```

	SUM_MOV_OBJ	ACTIVITY_TAG
1	1 car β€"	passenger
2	11 car β€"	driver

Εικόνα 23: Αποτελέσματα ερωτήματος

Το ίδιο ερώτημα με το προηγούμενο, αυτή τη φορά χωρίς την επιβολή του περιορισμού ότι το επόμενο STOP θα είναι «Working».

```

SELECT DEREf (tlink).o_id o_id,
       DEREf (tlink).traj_id traj_id,
       DEREf (tlink).sub_mpoint mpoint,
       DEREf (tlink).subtraj_id subtraj_id,
       value(t).mbb.getrectangle(4326),
       MDSYS.sdo_geom.sdo_centroid (VALUE (t).mbb.getrectangle (4326), 0.01
) mbr_centroid,
       activity_tag,
       defining_tag
FROM TABLE (std.stb_range_episodes_mbr (sem_episode ('STOP', NULL,
NULL, hermes.sem_mbb (hermes.sem_st_point (5.162558,51.12519,
tau_tll.d_timepoint_sec (2007, 1, 1, 7, 00, 00 ) ),
hermes.sem_st_point (5.180134, 51.140135, tau_tll.d_timepoint_sec
(2007, 1, 31, 7, 00, 00 ) ) ), NULL ), sem_episode ('STOP', NULL,
NULL, NULL, NULL ), NULL, 'SEM_INDX' ) ) t

```

Οπτικοποίηση

```

DECLARE
  stop_mbb hermes.sem_mbb := hermes.sem_mbb (hermes.sem_st_point
(5.162558, 51.12519, tau_tll.d_timepoint_sec (2007, 1, 1, 7, 00, 00 )
), hermes.sem_st_point (5.180134, 51.140135, tau_tll.d_timepoint_sec
(2007, 1, 31, 7, 00, 00 ) ) );
  res_mps mp_array := mp_array ();
  semtraj sem_trajectory;
  vsrid INTEGER := 4326;
BEGIN
  FOR rc IN
    (SELECT DEREf (tlink).o_id o_id,
           DEREf (tlink).traj_id traj_id,
           DEREf (tlink).sub_mpoint mpoint,

```



```

    Deref (tlink).subtraj_id subtraj_id,
    VALUE (t).mbb.getrectangle (vsrid) rect,
    MDSYS.sdo_geom.sdo_centroid (VALUE (t).mbb.getrectangle (vsrid),
    0.01 ) mbr_centroid,
    activity_tag,
    defining_tag
FROM TABLE (std.stb_range_episodes_mbr (sem_episode ('STOP', NULL,
NULL, stop_mbb, NULL ), sem_episode ('STOP', NULL, NULL, NULL, NULL ),
NULL, 'SEM_INDX' ) ) t
)
LOOP
    res_mps := mp_array (rc.mpoint);
    visualizer.movingpointtable2kml (res_mps, vsrid, CONCAT ( 'u' ||
rc.o_id || 'traj' || rc.traj_id || 'subtraj' || rc.subtraj_id,
'_Q_MOVPOINT.kml' ) );
    visualizer.polygon2kml (rc.rect, vsrid, rc.o_id || '.' ||
rc.traj_id || '.' || rc.subtraj_id || '_RECTANGLE.kml' );
    visualizer.placemark2kml (rc.mbr_centroid, vsrid, 'u' || rc.o_id
|| 'traj' || rc.traj_id || 'subtraj' || rc.subtraj_id ||
'_CENTROID.kml', ' ', rc.defining_tag || '-' || rc.activity_tag );
END LOOP;
    visualizer.polygon2kml (stop_mbb.getrectangle (vsrid), vsrid,
'RECTANGLE.kml' );
END;

```



Εικόνα 24: Στην εικόνα φαίνονται - με μπλε χρώμα - όλα τα επεισόδια με ετικέτα 'MOVE', τα οποία ξεκινούν από επεισόδιο με ετικέτα 'STOP' με δοθέν χωροχρονικό περιορισμό και καταλήγουν σε επεισόδιο με ετικέτα 'STOP'. Ο δοθέν χωρικός περιορισμός φαίνεται με πράσινο περίγραμμα.

5.3.10 Η αποθήκη δεδομένων σημασιολογικά επαυξημένων τροχιών

Για τη δημιουργία των αντικειμένων της αποθήκης δεδομένων και τον υπολογισμό των παραγόμενων δεδομένων ακολουθείται η εξής διαδικασία:

α) Δημιουργούμε τα αντικείμενα της αποθήκης δεδομένων δίνοντας επιθυμητό πρόθεμα.

```
begin
sdw.createsdw('sem_dw');
end;
```

β) Φορτώνουμε δεδομένα στους πίνακες διαστάσεων δίνοντας σαν ορίσματα το επιλεγμένο πρόθεμα που χρησιμοποιήσαμε κατά την δημιουργία των αντικειμένων, το όνομα του πίνακα με τα σημεία ενδιαφέροντος, το όνομα του πίνακα με τα χωροχρονικά μέγιστα-ελάχιστα του συνόλου δεδομένων μας, το όνομα του πίνακα με την πληροφορία χρηστών, το μέγεθος του χρονικού βήματος σε δευτερόλεπτα και το όνομα του πίνακα με τις σημασιολογικές τροχιές.

```
begin
sdw.loaddimensions('sem_dw','belg_pois','belg_dataset_dimensions','bel
g_users',2592000,'belg_sem_traj');
end;
```

γ) Φορτώνουμε δεδομένα στους πίνακες γεγονότων (STOPS-Fact-table, MOVES-Fact-table) δίνοντας το όνομα του πίνακα με τις σημασιολογικές τροχιές και το πρόθεμα του STB-tree ευρετηρίου που έχουμε δημιουργήσει σε αυτόν.

```
begin
sdw.cellstopsload('sem_dw','SEM_INDX');
end;
```

```
begin
sdw.cellmovesload('sem_dw','SEM_INDX');
end;
```

5.3.11 Εφαρμογή ερωτημάτων στην αποθήκη δεδομένων σημασιολογικά επαυξημένων τροχιών

Εφόσον έχει ολοκληρωθεί η ETL διαδικασία, έτσι όπως περιγράφηκε πιο πάνω, μπορούμε να εξετάσουμε διάφορα σενάρια στα δεδομένα αυτά.

Βρίσκουμε την συνολική μέση απόσταση, ανά φύλλο και ανά χρονική περίοδο όσων κινήθηκαν με τρένο.

```
SELECT sem_dw_period_dim.period_id, sem_dw_user_profile_dim.gender,
SUM (sem_dw_moves_fact.avg_travel_duration) sum_avg_dur,
```

```

SUM (sem_dw_moves_fact.avg_distance_traveled) sum_avg_dist
FROM      sem_dw_moves_fact,      sem_dw_user_profile_dim,
sem_dw_move_sems_dim, sem_dw_period_dim
WHERE      ((sem_dw_user_profile_dim.user_profile_id
=sem_dw_moves_fact.user_profile_id)
AND      (sem_dw_move_sems_dim.move_sems_id
=sem_dw_moves_fact.move_sems_id)
AND      (sem_dw_period_dim.period_id
=sem_dw_moves_fact.period_id)
AND      ((sem_dw_move_sems_dim.move_mode = 'TRAIN'))
) GROUP BY      sem_dw_user_profile_dim.gender,
sem_dw_period_dim.period_id ORDER BY sem_dw_period_dim.period_id ASC

```

PERIOD_ID	GENDER	SUM_AVG_DUR	SUM_AVG_DIST
1	1 MAN	5033	121811,56450774486
2	1 WOMAN	1783	24741,87326075254
3	2 MAN	1472	21255,3924289534
4	4 MAN	3357	89455,2482850577
5	4 WOMAN	647	8254,98104342971
6	5 MAN	923	11103,0588393071
7	6 MAN	5894	78477,9676656374
8	6 WOMAN	255	3343,68573034657
9	7 MAN	2064	23325,38109972392
10	7 WOMAN	455	4433,28462431386
11	8 WOMAN	640	3894,02341379328
12	10 MAN	337	6326,86237521639
13	10 (null)	676	6541,72758176443

Εικόνα 25: Εύρεση συνολικής διάρκειας και απόστασης ανά χρονική περίοδο και φύλο όσων κινήθηκαν με τρένο.

Βρίσκουμε τον αριθμό των χρηστών οι οποίοι κατευθύνονται σε ένα προορισμό και τον αριθμό των χρηστών που βρίσκονται ήδη στον ίδιο προορισμό μέσα στην ίδια χρονική περίοδο.

```
WITH period_stopsto AS
```

```

(SELECT      sem_dw_stops_fact.period_id,COUNT
(sem_dw_stops_fact.num_of_users)
num_of_stop_users,sem_dw_stop_sems_dim.stop_name
FROM      sem_dw_stops_fact,sem_dw_stop_sems_dim
WHERE      ((sem_dw_stop_sems_dim.stop_sems_id
=sem_dw_stops_fact.stop_sems_id ) )
GROUP BY      sem_dw_stops_fact.period_id,sem_dw_stop_sems_dim.stop_name
),
period_movesto AS

```

```

(SELECT
    sem_dw_MOVES_fact.period_id,COUNT
(sem_dw_MOVES_fact.num_of_users)
num_of_move_users,sem_dw_STOP_sems_dim.stop_name
FROM sem_dw_MOVES_fact,sem_dw_STOP_sems_dim
WHERE ((sem_dw_stop_sems_dim.stop_sems_id
sem_dw_MOVES_fact.to_stop_sems_id) )
GROUP BY sem_dw_MOVES_fact.period_id,sem_dw_STOP_sems_dim.stop_name
)
SELECT
period_stopsto.period_id,period_stopsto.stop_name,period_stopsto.num_o
f_stop_users,period_movesto.num_of_move_users
FROM period_stopsto FULL OUTER JOIN period_movesto
ON (period_stopsto.period_id=period_movesto.period_id AND
period_stopsto.stop_name=period_movesto.stop_name) ORDER BY 1 DESC,2;

```

PERIOD_ID	STOP_NAME	NUM_OF_STOP_USERS	NUM_OF_MOVE_USERS
25	philips	1	(null)
25	phillps	1	(null)
25	thuis	6	(null)
24	apothEEK	1	(null)
24	carpet	1	(null)
24	cc	1	(null)
24	cevo	1	1
24	dochter	1	(null)
24	hengelen	1	(null)
24	ipb	1	(null)
24	kantoor	1	1
24	kliniek	1	(null)
24	moe poes	2	(null)
24	opstaan en eten	1	(null)
24	philips	1	(null)
24	school	1	1

Εικόνα 26: Αριθμός κινούμενων αντικειμένων προς/σε προορισμό την ίδια χρονική περίοδο.

5.4 Εφαρμογή Επεκτάσεων

5.4.1 Απεικόνιση σημασιολογικά επαυξημένης τροχιάς.

Η διαδικασία *visualizer.semtrajectory2kml* οπτικοποιεί μια δοθείσα σημασιολογική τροχιά εμφανίζοντας τα MBRs, την σημασιολογική πληροφορία των επεισοδίων και τις αντίστοιχες υπο-τροχιές τους. Τα ορίσματα που δέχεται είναι η σημασιολογική τροχιά προς οπτικοποίηση, και τρεις boolean flags, οι οποίες ελέγχουν την εμφάνιση ή όχι των moving points, του MBR και του κεντροειδούς του κάθε επεισοδίου της τροχιάς.


```

DECLARE

semtraj      sem_trajectory;

void          INTEGER      := 123043;

vsemtrajid   INTEGER      := 1;

BEGIN

SELECT VALUE (s)

INTO semtraj

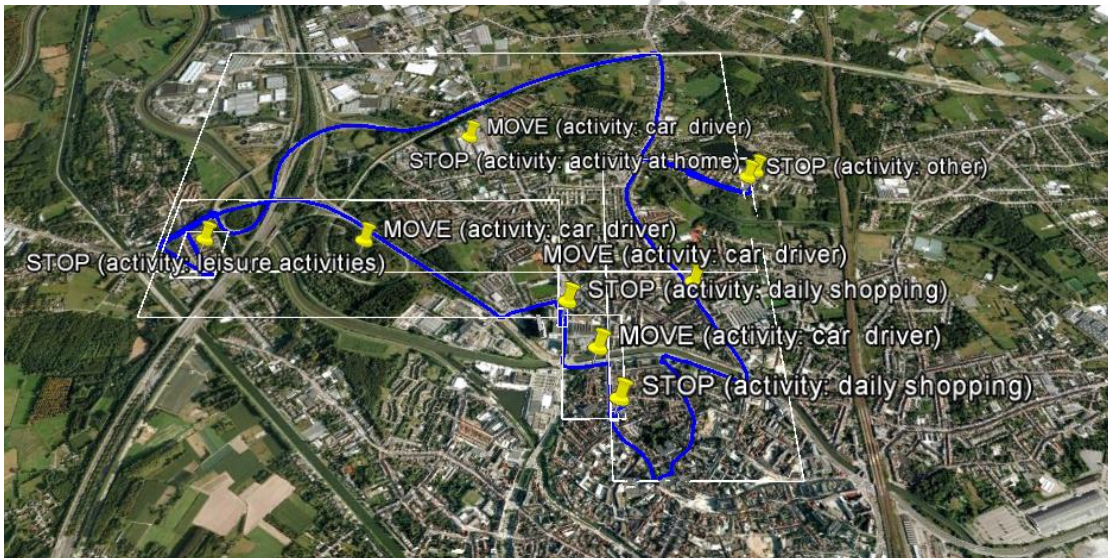
FROM belg_sem_trajs s

WHERE o_id = void AND semtraj_id = vsemtrajid;

visualizer.semtrajectory2kml (semtraj, 'TRUE', 'TRUE', 'TRUE');

END;

```



Εικόνα 27: Χρήση προσθήκης για την απεικόνιση σημασιολογικής τροχιάς. Φαίνονται τα MBRs των επεισοδίων με άσπρο περίγραμμα, η πληροφορία καθώς και οι υπο-τροχιές με μπλε χρώμα.

5.4.2 Εύρεση K πλησιέστερων σημείων ενδιαφέροντος.

Με τη χρήση της *sem_reconstruct.nn_pois*, βρίσκουμε τα K κοντινότερα σημεία ενδιαφέροντος από το κεντροειδές του ελάχιστου περιβάλλοντος ορθογωνίου (MBR) ενός επεισοδίου με παράλληλη απεικόνιση του αποτελέσματος. Τα ορίσματα που δέχεται είναι το MBR του επεισοδίου, ο αριθμός των K κοντινότερων σημείων που θέλουμε να βρούμε και μια boolean flag, η οποία ελέγχει την εμφάνιση ή όχι των αποτελεσμάτων (δηλ. την παραγωγή kml αρχείων).

```

DECLARE

sembb sem_mbb := sem_mbb (sem_st_point (5.503502, 50.953662,
tau_tll.d_timepoint_sec (2006, 5, 1, 5, 00, 00) ), sem_st_point

```

Διαχείριση σημασιολογικά επαυξημένων τροχιών κινούμενων αντικειμένων

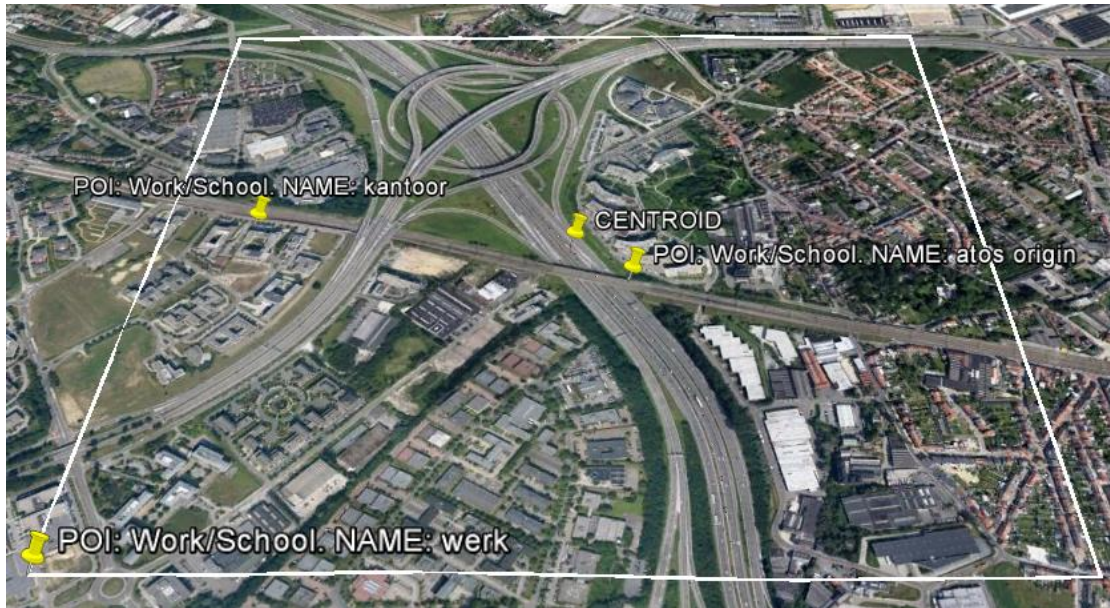
```
(5.727093, 51.026928, tau_tll.d_timepoint_sec (2007, 6, 1, 5, 00, 00)
) );
```

```
BEGIN
```

```
sem_reconstruct.nn_pois (sembb, 10, 'TRUE');
```

```
COMMIT;
```

```
END;
```



Εικόνα 28: Αποτελέσματα εύρεσης $K(=10)$ πλησιέστερων σημείων ενδιαφέροντος (POI), τα οποία βρίσκονται εντός του MBR (δείχνετε με άσπρο περίγραμμα) ενός επεισοδίου.

5.4.3 Εύρεση πιθανοτήτων για τα K πλησιέστερα σημεία ενδιαφέροντος ενός επεισοδίου.

Γίνεται χρήση της *sem_reconstruct.pois_probability* για εύρεση των πιθανοτήτων των K πλησιέστερων σημείων ενδιαφέροντος, τα οποία βρίσκονται εντός του MBR δοθέντος επεισοδίου. Έτσι προκύπτει το σημείο ενδιαφέροντος στο οποίο πραγματοποιήθηκε η δραστηριότητα με παράλληλη απεικόνιση του αποτελέσματος. Τα ορίσματα που δέχεται είναι το MBR του επεισοδίου και μια boolean flag, η οποία ελέγχει την εμφάνιση ή όχι των αποτελεσμάτων.

```
DECLARE
```

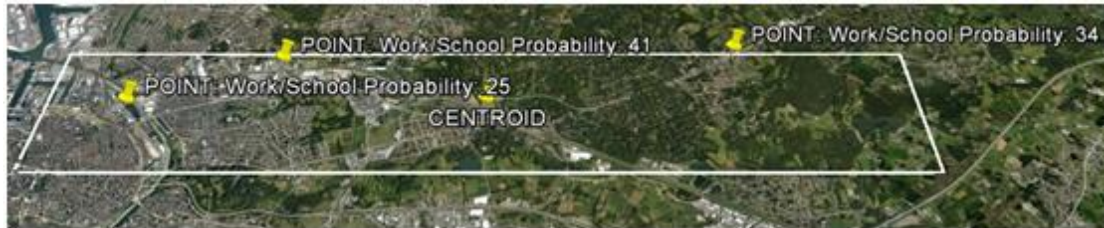
```
sembb sem_mbb := sem_mbb (hermes.sem_st_point (4.415422, 51.218345,
tau_tll.d_timepoint_sec (2007, 9, 4, 6, 27, 17) ),
hermes.sem_st_point (4.634043, 51.24343, tau_tll.d_timepoint_sec
(2007, 9, 5, 5, 44, 37) ) );
```

```
BEGIN
```

```
sem_reconstruct.pois_probability (sembb, 'TRUE');
```

```
COMMIT;
```

```
END;
```



Εικόνα 29: Αποτελέσματα εύρεσης των πιθανοτήτων των $K(=10)$ πλησιέστερων σημείων ενδιαφέροντος (POI), τα οποία βρίσκονται εντός του MBR (δείχνετε με άσπρο περίγραμμα) ενός επεισοδίου.

5.4.4 Δημιουργία μήτρας προέλευσης-προορισμού.

Η δημιουργία μιας μήτρας προέλευσης-προορισμού από σημασιολογικά επανυημένες τροχιές, όπως περιγράφεται στον αλγόριθμο Εικόνα 8: Αλγόριθμος `construct_odmatrix`, απαιτεί τον διαχωρισμό σε κελία της περιοχής που οριοθετούν οι τροχιές, ώστε να υπολογιστεί το τελικό αποτέλεσμα. Για το λόγο αυτό κατά την χρήση της `od_matrix.populate_rectangle_tbl (stepx, stepy)` δίνουμε το επιθυμητό βήμα κατά γεωγραφικό πλάτος και μήκος.

```

DECLARE
    stepx  BINARY_INTEGER;
    stepy  BINARY_INTEGER;
BEGIN
    stepx := 0.5;
    stepy := 0.5;
    hermes.od_matrix.populate_rectangle_tbl (stepx, stepy);
    COMMIT;
END;

```

Στη συνέχεια μπορούμε να πάρουμε τη παραγόμενη μήτρα με χρήση της `od_matrix.get_odmatrix`.

```

SELECT *
FROM TABLE (hermes.od_matrix.get_odmatrix)
;

```


	STARTCELL	ENDCELL	TOTAL
1	859692	859692	95
2	859689	859689	70
3	859691	859691	46
4	859688	859688	25
5	859686	859686	12
6	859685	859685	11
7	859695	859695	5
8	859689	859688	2
9	859688	859689	2
10	859695	859691	1
11	859692	859691	1
12	859692	859688	1
13	859691	859694	1
14	859691	859692	1
15	859691	859689	1
16	859689	859686	1

Εικόνα 30: Η παραγόμενη μήτρα προέλευσης-προορισμού από σημασιολογικά επαυξημένες τροχιές.

Η απαίτηση τμηματοποίησης του χώρου υπάρχει και στην περίπτωση της μήτρας όταν υπολογίζεται από την αποθήκη δεδομένων, γι αυτό και το πρώτο βήμα είναι πανομοιότυπο. Ακολουθεί η κλήση της αντίστοιχης διαδικασίας *od_matrix.get_dw_odmatrix* κατά την οποία δίνεται σαν όρισμα το πρόθεμα το οποίο δηλώσαμε κατά την δημιουργία των αντικειμένων της αποθήκης δεδομένων. Ο υπολογισμός των αποτελεσμάτων βασίζεται στον πίνακα γεγονότων των MOVE επεισοδίων.

```

DECLARE
    stepx  BINARY_INTEGER;
    stepy  BINARY_INTEGER;
BEGIN
    stepx := 0.5;
    stepy := 0.5;
    hermes.od_matrix.populate_rectangle_tbl (stepx, stepy);
    COMMIT;
END;

SELECT *
FROM TABLE (hermes.od_matrix.get_dw_odmatrix('SEM_DW')) ;

```

Και στις δύο περιπτώσεις η μήτρα εμφανίζεται με τη μορφή σχεσιακού πίνακα, όπου οι δύο πρώτες στήλες δηλώνουν το αναγνωριστικό των κελιών, ενώ η τρίτη στήλη το τελικό αποτέλεσμα. Για να είναι όμως δυνατή η σύγκριση των αποτελεσμάτων υλοποιήθηκε και μια διαδικασία μετασχηματισμού, όπως περιγράφεται στην παράγραφο 4.3.4, κάθε τροχιάς σε πολλές τροχιές με βάση τα STOP και MOVE επεισόδια της. Με τον τρόπο αυτό εφαρμόζοντας τον πρώτο αλγόριθμο στον εξαγόμενο πίνακα σημασιολογικών τροχιών που δημιουργεί ο μετασχηματισμός, αναμένουμε τα ίδια αποτελέσματα με αυτά του δεύτερου αλγορίθμου μήτρας προέλευσης-προορισμού, ο οποίος εφαρμόζεται στον πίνακα γεγονότων των MOVE επεισοδίων. Η χρήση της για μετασχηματισμό των σημασιολογικά επαυξημένων τροχιών του πίνακα 'BELG_SEM_TRAJS' γίνεται ως εξής:

```
BEGIN
```

```
HERMES.MAPPING_SEM_TRAJ.MAPPING_SEMTRAJS ( 'BELG_SEM_TRAJ' );
```

```
END;
```

Πανεπιστήμιο Πειραιώς

6 Συμπεράσματα

Καθώς η έρευνα σε βάσεις δεδομένων σημασιολογικά επαυξημένων τροχιών βρίσκεται ακόμη σε αρχικό στάδιο, τα ζητήματα που πρέπει να αντιμετωπιστούν είναι πολλά. Η επέκταση του HERMES MOD αποτελεί ένα πρότυπο πλαίσιο προς την κατεύθυνση αυτή. Με την παρούσα εργασία ανακαλύψαμε τις δυνατότητες τόσο του πλαισίου αυτού όσο και τις δυνατότητες εκμετάλλευσης των σημασιολογικά επαυξημένων τροχιών. Είδαμε την “ευκολία” παραγωγής γνώσης από δεδομένα αυτού του τύπου, εφαρμόζοντας υπάρχοντες τελεστές για την εκμετάλλευση τόσο της χωροχρονικής πληροφορίας των δεδομένων όσο και της σημασιολογικής.

Και ενώ ένα μεγάλο τμήμα της διαχείρισης των σημασιολογικά επαυξημένων τροχιών υποστηρίζεται από το HERMES MOD, υπάρχουν ανοιχτά ζητήματα που πρέπει να αντιμετωπιστούν. Ως προς την κατεύθυνση αυτή αναπτύχθηκαν και οι νέες προθήκες του συστήματος, συνεισφέροντας σε κύρια τμήματα, όπως είναι η οπτικοποίηση των αποτελεσμάτων, η ανασύσταση των τροχιών και η εξαγωγή γνώσης.

Πιστεύουμε πως οι αρχικοί στόχοι της εργασίας επετεύχθησαν και πως έγινε μια αν και περιορισμένη, ουσιαστική παρουσίαση του αντικειμένου.

Πανεπιστήμιο Πειραιώς

7 Βιβλιογραφία

- CHRISTINE PARENT, S. S.-D. (2011). *Semantic Trajectories Modeling and Analysis*.
- G. Agamennoni, J. N. (2009). Mining GPS Data for Extracting Significant Places. *ICRA*.
- Han, J. (1995). *Mining Knowledge at Multiple Concept Levels (An invited talk)*.
- L. O. Alvares, V. B. (2007). *Towards Semantic Trajectory*. Belgium: Hasselt University.
- M. Zimmermann, T. K. (2009). Finding Stops in Error-Prone Trajectories of Moving Objects with Time-Based Clustering. Στο *Intelligent Interactive Assistance and Mobile Multimedia Computing* (σσ. 275–286).
- N. Pelekis. (2002). *STAU: A spatio-temporal extension for the ORACLE DBMS*. UMIST.
- Nikos Pelekis, E. F. (2011). *HERMES: A Trajectory DB Engine for Mobility-Centric Applications*. International Journal of Knowledge-based Organizations.
- R.H. Güting, M. B. (2000). A Foundation for Representing and Querying Moving Objects. *ACM Transactions on Database System* 25 , σσ. 1–42.
- R.H. Güting, T. B. (2004). *SECONDO: An Extensible DBMS Architecture and Prototype*. Hagen: Fernuniversität.
- R.H. Güting, T. B. (2004). *SECONDO: An Extensible DBMS Architecture and Prototype*. *Informatik Report 313*. Hagen.
- S. Vosinakis N. Pelekis, Y. T. (2006). Hermes - a framework for location-based data management. *In Proc. EDBT*.
- Sotiris Brakatsoulas, D. P. (2004). Modeling, Storing and Mining Moving Object Databases. *Database Engineering and Applications Symposium, 2004. IDEAS '04. Proceedings. International*.
- Spaccapietra S, P. C. (2008). A conceptual view on trajectories. Στο *Data and Knowledge Engineering* (σσ. 126-146).
- SPACCAPIETRA, S. A. (2011). Adding meaning to your steps. *Proc. of the ER 2011 Int. Conf.* Brussels.
- Starner, D. A. (2003). Using GPS to Learn Significant Locations. Στο *Personal and Ubiquitous Computing* (σσ. 275–286).
- Vania Bogorny, C. A. (2010). A Conceptual Data Model for Trajectory Data Mining. Στο *A Conceptual Data Model for Trajectory Data Mining* (σσ. 1-15). Zurich: Springer Berlin Heidelberg.
- Vania Bogorny, C. R. (2013). CONSTAnT - A Conceptual Data Model for Semantic Trajectories of Moving Objects. *Transaction in GIS*.
- Yan, Z. (2011). *Semantic Trajectories: Computing and Understanding Mobility Data*. Lausanne: Swiss Federal Institute of Technology in Lausanne.
- Z. Yan, C. P. (2010). A Hybrid Model and Computing Platform for Spatio-Semantic Trajectories. *ESWC*.
- ZHIXIAN YAN, D. C. (2012). *Semantic Trajectories: Mobility Data Computation and Annotation*.
- Zhixian Yan, D. C. (2011). SeMiTri: A Framework for Semantic Annotation of Heterogeneous Trajectories. *EDBT:259--270*.
- Zhixian Yan, N. G. (2011). SeTraStream: Semantic-Aware Trajectory Construction over Streaming Movement Data. *the 12th Intl. Symposium on Spatial and Temporal Databases (SSTD2011)*. Minneapolis, MN, USA.
- Μανταδάκης Ευστάθιος, Μ. Γ. (2008). *ΑΝΑΛΥΣΗ ΚΑΙ ΟΠΤΙΚΟΠΟΙΗΣΗ ΤΡΟΧΙΩΝ ΚΙΝΟΥΜΕΝΩΝ ΑΝΤΙΚΕΙΜΕΝΩΝ – ΤΟ ΣΥΣΤΗΜΑ TRAJECTORY SURFER*. Πειραιάς: Πανεπιστήμιο Πειραιώς.