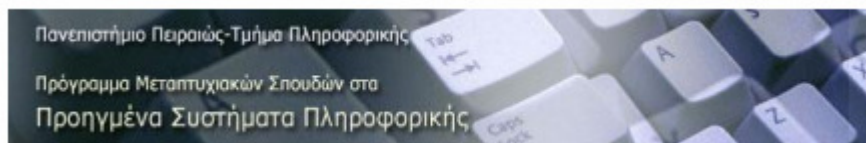




Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

| | |
|-----------------------|---|
| Τίτλος Διατριβής | Εύρεση Πιθανής Περιοχής Δραστηριοποίησης Κινούμενων Αντικειμένων |
| Όνοματεπώνυμο Φοιτητή | Μούσο Γιουρεντίνα |
| Αριθμός Μητρώου | ΜΠΣΠ / 11036 |
| Κατεύθυνση | Συστήματα Υποστήριξης Αποφάσεων |
| Επιβλέπων | Νίκος Πελέκης, Λέκτορας |



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

(υπογραφή)

(υπογραφή)

Νικόλαος Πελέκης
Λέκτορας

Ιωάννης Θεοδωρίδης
Καθηγητής

Ιωάννης Σίσκος
Καθηγητής

ΠΕΡΙΕΧΟΜΕΝΑ

| | |
|---|----|
| <i>ΠΕΡΙΕΧΟΜΕΝΑ</i> | 4 |
| <i>ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ</i> | 5 |
| <i>ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ</i> | 6 |
| <i>ΠΕΡΙΛΗΨΗ</i> | 8 |
| <i>ABSTRACT</i> | 9 |
| ΚΕΦΑΛΑΙΟ 1 | 10 |
| Εισαγωγή | 10 |
| 1.1 Σύντομη Περιγραφή | 11 |
| 1.2 Δομή Εργασίας..... | 12 |
| ΚΕΦΑΛΑΙΟ 2 | 14 |
| Σχετικές Εργασίες..... | 14 |
| 2.1 Εισαγωγή | 15 |
| 2.2 Χωρικά και Χωροχρονικά Δεδομένα..... | 17 |
| 2.3 Χωροχρονικές Βάσεις Δεδομένων | 17 |
| 2.4 Κινούμενα Αντικείμενα | 19 |
| 2.5 Τροχιά Κινούμενων Αντικειμένων..... | 20 |
| 2.6 Πιθανή Περιοχή Δραστηριοποίησης Κινούμενων Αντικειμένων | 22 |
| 2.7 Σημεία Ενδιαφέροντος σε Περιοχή Δραστηριοποίησης εν Κινήσει Χρηστών | 25 |
| 2.8 Ανακάλυψη Δημοφιλής Διαδρομής από Ιστορικές Τροχιές..... | 29 |
| 2.9 Μείωση Αβεβαιότητας Χαμηλού Ρυθμού Δειγματοληψίας σε Τροχιές | 34 |
| 2.10 Map Matching | 39 |
| ΚΕΦΑΛΑΙΟ 3 | 48 |
| Εύρεση Πιθανής Περιοχής Δραστηριοποίησης..... | 48 |
| 3.1 Εισαγωγή | 49 |
| 3.2 Εργαλεία | 49 |
| 3.2.1 Η PostgreSQL | 49 |
| 3.2.2 Το PostGIS | 50 |
| 3.2.3 Το Quantum GIS | 51 |
| 3.3 Potential Area of Activity (PAA) | 51 |
| 3.4 Περιγραφή Υλοποίησης της PAA | 54 |
| 3.5 Οπτικοποίηση Αποτελεσμάτων | 56 |
| 3.5.1 Παράδειγμα 1 | 56 |
| 3.5.2 Παράδειγμα 2 | 58 |
| ΚΕΦΑΛΑΙΟ 4 | 61 |
| Το Map Matching | 61 |
| 4.1 Εισαγωγή | 62 |
| 4.2 Εργαλεία | 62 |
| 4.2.1 Το pgRouting | 62 |
| 4.2.2 Το Osm2po | 62 |
| 4.3 Περιγραφή Map Matching Συναρτήσεων | 63 |

| | |
|--|----|
| 4.4 Περιγραφή Υλοποίησης του Map Matching..... | 67 |
| 4.5 Οπτικοποίηση Αποτελεσμάτων | 71 |
| ΣΥΜΠΕΡΑΣΜΑΤΑ | 77 |
| ΒΙΒΛΙΟΓΡΑΦΙΑ | 78 |

ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ

| | |
|--|----|
| Εικόνα 2-1: Αναπαράσταση Δεδομένων | 17 |
| Εικόνα 2-2: Σχέση μεταξύ χωρικών, χρονικών και χωροχρονικών DBMS..... | 18 |
| Εικόνα 2-3: Τροχιά ενός κινούμενου αντικείμενου | 21 |
| Εικόνα 2-4: Τροχιές στο χωροχρονικό διάστημα | 21 |
| Εικόνα 2-5: Πιθανές τροχιές κινούμενου αντικείμενου..... | 22 |
| Εικόνα 2-6: Η τομή των δύο κύκλων | 23 |
| Εικόνα 2-7: Έλλειψη..... | 24 |
| Εικόνα 2-8: Ρυθμός δειγματοληψίας..... | 24 |
| Εικόνα 2-9: Εννοιολογικό πλαίσιο..... | 26 |
| Εικόνα 2-10: Γενική δομή του σχεδιασμού | 27 |
| Εικόνα 2-11: Η γραφική διεπαφή χρήστη | 28 |
| Εικόνα 2-12: Απεικόνιση εστιατορίων της πόλης της Γάνδης στο Google_Earth..... | 29 |
| Εικόνα 2-13: Παραδείγματα ανακάλυψης δημοφιλών διαδρομών..... | 30 |
| Εικόνα 2-14: Εξόρυξη Δικτύου Μεταφοράς..... | 31 |
| Εικόνα 2-15: Διανομή Πιθανότητας Μεταφοράς..... | 32 |
| Εικόνα 2-16: Maximum Probability Product Αλγόριθμος..... | 33 |
| Εικόνα 2-17: Παραδείγματα Ερωτημάτων (A-Start Node, B-End Node)..... | 33 |
| Εικόνα 2-18: Επισκόπηση Συστήματος..... | 35 |
| Εικόνα 2-19: Τροχιά αναφοράς..... | 36 |
| Εικόνα 2-20: Κατασκευή γραφήματος διάσχισης..... | 37 |
| Εικόνα 2-21: Nearest neighbour based inference | 37 |
| Εικόνα 2-22: Global route inference..... | 38 |
| Εικόνα 2-23: Το Map-Matching Πρόβλημα..... | 39 |
| Εικόνα 2-24: Το Map-Matching..... | 40 |
| Εικόνα 2-25: Ένα πρόβλημα με point-to-point matching..... | 41 |
| Εικόνα 2-26: Εφαρμογή καμπύλης..... | 41 |
| Εικόνα 2-27: Η απόσταση μεταξύ ενός σημείου και ενός τμήματος | 42 |
| Εικόνα 2-28: Ένα πρόβλημα αντιστοίχισης σημείου-προς-καμπύλη | 42 |
| Εικόνα 2-29: Ένα άλλο πρόβλημα αντιστοίχισης σημείου-προς-καμπύλη..... | 43 |
| Εικόνα 2-30: Χρησιμοποιώντας τοπολογικές πληροφορίες..... | 43 |
| Εικόνα 2-31: Αποστάσεις Καμπύλη-προς-καμπύλη | 44 |
| Εικόνα 2-32: Η αρχή του Map-Matching | 45 |
| Εικόνα 2-33: Επισκόπηση της αρχιτεκτονικής του συστήματος..... | 46 |
| Εικόνα 3-1: pgAdmin, το περιβάλλον διεπαφής της PostgreSQL..... | 50 |
| Εικόνα 3-2: Ψευδοκώδικας της συνάρτησης PAA..... | 53 |
| Εικόνα 3-3: Τα σημεία της έλλειψης | 54 |
| Εικόνα 3-4: Παράδειγμα_1_ Περίπτωση_1 | 57 |
| Εικόνα 3-5: Παράδειγμα_1_ Περίπτωση_2 | 57 |

| | |
|--|----|
| Εικόνα 3-6: Παράδειγμα_1_ Περίπτωση_3..... | 58 |
| Εικόνα 3-7: Παράδειγμα_2_ Περίπτωση_1..... | 59 |
| Εικόνα 3-8: Παράδειγμα_2_ Περίπτωση_2..... | 59 |
| Εικόνα 3-9: Παράδειγμα_2_ Περίπτωση_3..... | 60 |
| Εικόνα 4-1: Ο πίνακας του δικτύου milano_network..... | 68 |
| Εικόνα 4-2: Ο πίνακας με τα χωροχρονικά δεδομένα..... | 69 |
| Εικόνα 4-3: Ο πίνακας με τα map-matched αποτελέσματα..... | 70 |
| Εικόνα 4-4: Ο πίνακας του δικτύου του Μιλάνου..... | 71 |
| Εικόνα 4-5: Παράδειγμα 1..... | 72 |
| Εικόνα 4-6: Το map-matching αποτέλεσμα του παραδείγματος 1..... | 72 |
| Εικόνα 4-7: Παράδειγμα 2..... | 73 |
| Εικόνα 4-8: Το map-matching αποτέλεσμα του παραδείγματος 2..... | 73 |
| Εικόνα 4-9: Παράδειγμα 3..... | 74 |
| Εικόνα 4-10: Το map-matching αποτέλεσμα του παραδείγματος 3..... | 74 |
| Εικόνα 4-11: Παράδειγμα 4..... | 75 |
| Εικόνα 4-12: Το map-matching αποτέλεσμα του παραδείγματος 4..... | 75 |
| Εικόνα 4-13: Παράδειγμα 5..... | 76 |
| Εικόνα 4-14: Το map-matching αποτέλεσμα του παραδείγματος 5..... | 76 |

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ

| | |
|--|----|
| Πίνακας 2-1: Παράμετροι της συνάρτησης πιθανότητας P2..... | 23 |
| Πίνακας 3-1: Περιγραφή Δεδομένων..... | 54 |
| Πίνακας 3-2: Παράμετροι της συνάρτησης PAA..... | 55 |
| Πίνακας 3-3: Συναρτήσεις του PostGIS που χρησιμοποιήθηκαν..... | 56 |
| Πίνακας 4-1: Περιγραφή Χωροχρονικών Δεδομένων..... | 68 |
| Πίνακας 4-2: Ο πίνακας με τα map-matched αποτελέσματα..... | 69 |
| Πίνακας 4-3: Συναρτήσεις του PostGIS που χρησιμοποιήθηκαν..... | 70 |

Ευχαριστίες

Στο σημείο αυτό, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Νίκο Πελέκη για τη δυνατότητα που μου έδωσε να ασχοληθώ με την παρούσα μεταπτυχιακή εργασία.

Επίσης, ιδιαίτερες ευχαριστίες, οφείλω στον υποψήφιο διδάκτορα κ. Παναγιώτη Ταμπάκη για την βοήθεια που μου παρείχε για τις όποιες δυσκολίες αντιμετώπισα σε όλη τη διάρκεια της παρούσας εργασίας.

Τέλος να ευχαριστήσω παρά πολύ την οικογένειά μου για την στήριξη που μου έχει προσφέρει στη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Το αντικείμενο της παρούσας μεταπτυχιακής διατριβής είναι η μελέτη εύρεσης πιθανής περιοχής δραστηριοποίησης κινούμενων αντικειμένων. Το πρόβλημα του προσδιορισμού της πιθανής περιοχής δραστηριότητας ενός κινούμενου αντικείμενου βρίσκεται στην προσπάθεια να προσδιορίσει την περιοχή στην οποία το κινούμενο αντικείμενο δυνητικά θα κινηθεί σε μια δεδομένη χρονική περίοδο και με μία δεδομένη μέγιστη ταχύτητα. Δίνοντας ένα αντικείμενο, ένα χρονικό ορόσημο και την μέγιστη ταχύτητα του αντικειμένου, μπορούμε να υπολογίσουμε μια περιοχή γύρω από το αντικείμενο που θα είναι η περιοχή που ενδεχομένως θα κινηθεί.

Δημιουργήθηκε η συνάρτηση εύρεσης πιθανής περιοχής δραστηριότητας PAA του κινούμενου αντικειμένου. Η συνάρτηση PAA, η οποία βρίσκει τη πιθανή περιοχή δραστηριότητας ενός κινούμενου αντικείμενου δέχεται σαν είσοδο την μέγιστη ταχύτητα που μπορεί να κινείται το κινούμενο αντικείμενο, το id της τροχιάς, την ενδιάμεση χρονική στιγμή tx στο χρονικό διάστημα κίνησής του και την ενδιάμεση χρονική στιγμή ty στο χρονικό διάστημα κίνησής του.

Ένα άλλο πρόβλημα που μελετήθηκε στα πλαίσια της εργασίας είναι το map-matching πρόβλημα. Το map-matching είναι μια διαδικασία που έχει στόχο να προσεγγίζει όσο πιο πιστά γίνεται την τροχιά ενός κινούμενου αντικειμένου πάνω σε έναν ψηφιακό χάρτη και είναι ένα θεμελιώδες στάδιο προεπεξεργασίας για πολλές εφαρμογές, όπως η διαχείριση κινούμενων αντικειμένων.

Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα με πολλές δυνατότητες, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση PostGIS. Το PostGIS είναι ένα ανοικτό πρόγραμμα λογισμικού που προσθέτει υποστήριξη για γεωγραφικά αντικείμενα στην αντικείμενο-σχεσιακή βάση δεδομένων PostgreSQL. Επίσης, έχει χρησιμοποιηθεί το pgRouting που είναι επέκταση του PostGIS και της PostgreSQL και δίνει δυνατότητες δρομολόγησης σε γεωγραφικά δεδομένα και άλλες λειτουργίες ανάλυσης του δικτύου. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα QGIS.

ABSTRACT

The object of this thesis is the study of the discovery of potential area of activity of moving objects. The problem of determining of the potential area of activity of a moving object is to try to identify the area in which the moving object will potentially move in a given period and a given maximum speed. Giving an object a timestamp and the maximum speed of the object, we can compute a region around the object that is the area that might move.

Created the function of the discovery of potential area of activity of moving object, PAA. The function PAA, which finds the potential area of activity of a moving object accepts as input the maximum speed that can move the moving object, the id of the trajectory, the intermediate time t_x at the period of its movement and the intermediate time t_y at the period of its movement.

Another problem studied in this thesis is the map-matching problem. The map-matching is a process that aims to approximate as closely as possible the trajectory of a moving object on a digital map and is an essential pre-processing step for many applications, such as moving objects management.

For the implementation have been used the relational open source database with many features, PostgreSQL, enriched with spatial extension PostGIS. The PostGIS is an open source software that adds support for geographic objects in object - relational database PostgreSQL. It has also been used pgRouting which is an extension of PostGIS and PostgreSQL and provides routing capabilities to geographic data and other network analysis functions. To visualize the results used the program of QGIS.

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

1.1 Σύντομη Περιγραφή

Η παρούσα μεταπτυχιακή διατριβή εκπονείται στα πλαίσια του Προγράμματος Σπουδών Π.Μ.Σ «Προηγμένα Συστήματα Πληροφορικής» στο Τμήμα Πληροφορικής στο Πανεπιστήμιο Πειραιά.

Το αντικείμενο της παρούσας διατριβής είναι η μελέτη εύρεσης πιθανής περιοχής δραστηριοποίησης εν κινήσει χρηστών ή γενικότερα κινούμενου αντικείμενου. Εστιάζει στην χώρο-χρονική μελέτη της συνεχούς κίνησης ανθρώπου (ή γενικότερα κινούμενου αντικείμενου) δεδομένης της τροχιάς που έχει ακολουθήσει το αντικείμενο. Το αντικείμενο κινείται με μεταβλητή ταχύτητα, είναι γνωστή δε η μέγιστη ταχύτητα που μπορεί να επιτύχει.

Για την εκπόνηση της παρούσας διπλωματικής εργασίας χρειάστηκε αρχικά να μελετήσω μια σειρά από επιστημονικά άρθρα πάνω στα κινούμενα αντικείμενα και γενικότερα στα Γεωγραφικά Συστήματα Πληροφοριών (Geographic Information System). Αναφέρονται και περιγράφονται έννοιες, των χωρικών και χωροχρονικών δεδομένων, χωροχρονικών βάσεων δεδομένων, κινούμενων αντικείμενων, τροχιάς κίνησης, τροχιές κινούμενων αντικείμενων, η πιθανή περιοχή δραστηριοποίησης κινούμενων αντικείμενων. Τα χωρικά δεδομένα (spatial data), προσδιορίζουν τα γεωμετρικά χαρακτηριστικά του στοιχείου (θέση, διαστάσεις, σχήμα κ.λπ.) και έχουν άμεση σχέση με τον εντοπισμό του. Τα χωροχρονικά δεδομένα είναι τα δεδομένα που συνδυάζουν χωρική πληροφορία με τη διάσταση του χρόνου. Τα χωροχρονικά δεδομένα είναι μια κατηγορία δεδομένων που περιλαμβάνουν τόσο το χώρο όσο και το χρόνο σαν μεταβλητές [ΘΠ2012].

Επίσης, για την υλοποίηση της εργασίας ήταν απαραίτητη η μελέτη των χωροχρονικών βάσεων δεδομένων αφού τα τελευταία χρόνια έχουν γίνει πολύ σημαντικές, καθώς πολλές εφαρμογές πρέπει να αποθηκεύουν δεδομένα του πραγματικού κόσμου που δείχνουν χωρικά καθώς και χρονικά χαρακτηριστικά. Μια χωροχρονική βάση δεδομένων είναι μια βάση δεδομένων που διαχειρίζεται χωρικές και χρονικές πληροφορίες. Χωροχρονικές εφαρμογές που περιλαμβάνουν συνεχώς κινούμενα αντικείμενα είναι όλο και περισσότερο στο επίκεντρο της έρευνας. Έρευνα σε κινούμενα αντικείμενα έχει δημοσιευθεί στις εργασίες [GBEJ2000], [PJ1999]. Τα συστήματα χωροχρονικών βάσεων δεδομένων αποτελούν στην ουσία συστήματα κινούμενων αντικείμενων (moving objects) με ιδιαίτερο και ερευνητικά γόνιμο ενδιαφέρον, προσφέροντας τη δυνατότητα να συγκεραστούν ιδέες από τα συγγενικά πεδία των χωρικών και των χρονικών βάσεων δεδομένων. Ως χαρακτηριστικό παράδειγμα θα μπορούσε να χρησιμεύσει ένα υποθετικό σύστημα διευκόλυνσης της κυκλοφορίας των οχημάτων στο οδικό δίκτυο μιας πυκνοκατοικημένης περιοχής. Μια εκτεταμένη έρευνα χωροχρονικών βάσεων δεδομένων γίνεται στην εργασία [AR1999].

Στη συνέχεια για την εύρεση της πιθανής περιοχής δραστηριοποίησης κινούμενων αντικείμενων μελετήθηκαν αναλυτικά οι εργασίες [PJ1999], [NVS2010]. Το πρόβλημα του προσδιορισμού της πιθανής περιοχής δραστηριότητας ενός κινούμενου αντικείμενου βρίσκεται στην προσπάθεια να προσδιορίσει την περιοχή στην οποία το κινούμενο αντικείμενο δυνητικά θα κινηθεί σε μια δεδομένη χρονική περίοδο και με μία δεδομένη μέγιστη ταχύτητα. Δίνοντας ένα αντικείμενο, ένα χρονικό ορόσημο (timestamp) και την μέγιστη ταχύτητα του αντικείμενου, μπορούμε να υπολογίσουμε μια περιοχή γύρω από το αντικείμενο που θα είναι η περιοχή που ενδεχομένως θα κινηθεί. Αναλυτική περιγραφή για τις συγκεκριμένες εργασίες έχει γίνει στο δεύτερο κεφάλαιο όπου αναφέρεται όλο το θεωρητικό υπόβαθρο για την επίλυση του συγκεκριμένου προβλήματος. Σύμφωνα με την θεωρητική μελέτη που έχει παρουσιαστεί στις εργασίες αυτές υλοποιήθηκε η λύση του συγκεκριμένου προβλήματος.

Δημιουργήθηκε η συνάρτηση εύρεσης πιθανής περιοχής δραστηριότητας PAA (Potential Area of Activity) του κινούμενου αντικείμενου. Η συνάρτηση PAA, η οποία βρίσκει τη πιθανή περιοχή δραστηριότητας ενός κινούμενου αντικείμενου δέχεται σαν είσοδο την μέγιστη ταχύτητα που μπορεί να κινείται το κινούμενο αντικείμενο, το id της τροχιάς, την ενδιάμεση χρονική στιγμή tx στο χρονικό διάστημα κίνησης του και την ενδιάμεση χρονική στιγμή ty στο χρονικό διάστημα κίνησης του. Η αναλυτική περιγραφή της συνάρτησης εύρεσης πιθανής περιοχής δραστηριότητας PAA και ο ψευδοκώδικας δίνονται στο τρίτο κεφάλαιο της εργασίας. Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα με πολλές δυνατότητες, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση PostGIS. Το PostGIS είναι ένα ανοικτό πρόγραμμα λογισμικού που προσθέτει υποστήριξη για γεωγραφικά αντικείμενα στην αντικείμενο-σχεσιακή βάση δεδομένων PostgreSQL. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα QGIS.

Ένα άλλο πρόβλημα που μελετήθηκε στα πλαίσια της εργασίας είναι το map-matching πρόβλημα. Το map-matching είναι μια διαδικασία που έχει στόχο να προσεγγίζει όσο πιο πιστά γίνεται την τροχιά ενός κινούμενου αντικείμενου πάνω σε έναν ψηφιακό χάρτη. Δηλαδή γίνεται προσπάθεια να «ταιριάξουμε» την εκτιμωμένη θέση με μια καμπύλη στο χάρτη για να εκτιμηθεί με μεγαλύτερη ακρίβεια η πραγματική θέση του χρήστη. Με την βοήθεια της διαδικασίας map-matching μεταφέρουμε τις εκτιμήσεις στο οδικό δίκτυο και ως αποτέλεσμα έχουμε μια καλύτερη εικόνα της τροχιάς που ακολουθήθηκε. Άρα, το map-matching είναι η διαδικασία ευθυγράμμισης μιας ακολουθίας παρατηρούμενων θέσεων χρήστη με το οδικό δίκτυο σε έναν ψηφιακό χάρτη. Είναι ένα θεμελιώδες στάδιο προεπεξεργασίας για πολλές εφαρμογές, όπως η διαχείριση κινούμενου αντικείμενου, ανάλυση της ροής κυκλοφορίας. Υπάρχει ένας αριθμός από διαφορετικούς τρόπους για να προσεγγίσει το map matching πρόβλημα, καθένας από τους οποίους έχει πλεονεκτήματα και μειονεκτήματα. Μελέτη πάνω στο συγκεκριμένο πρόβλημα έγινε στις εργασίες [BK1996], [WBK2000], [ZG2006], [LZZX2009], [AWB2011].

Στην εργασία [WBK2000] προσεγγίζεται το map matching ως πρόβλημα αναζήτησης και ως στατιστική εκτίμηση. Ο συνδυασμός της απλότητας της απλής προσέγγισης αναζήτησης με ορισμένες από τις ιδέες της στατιστικής προσέγγισης εφαρμόζεται για τέσσερις διαφορετικούς αλγόριθμους. Οι τύποι δεδομένων που χειρίζεται το map matching περιλαμβάνουν point-to-line, line-to-line και polyline-to-polyline. Με βάση τα διαχρονικά χαρακτηριστικά, οι αλγόριθμοι map-matching μπορούν επίσης να ταξινομηθούν σε on-line map-matching και off-line map-matching [ZG2006]. Στην εργασία [AWB2011] μελετώνται τα ζητήματα που αφορούν τις map-matching αβεβαιότητες θέσης και αναλύονται οι παράγοντες που επηρεάζουν το map-matching. Στην εργασία αυτή, συζητήθηκαν οι τεχνολογίες του map matching και οι τεχνικές για το πρόβλημα της χαμηλής ακρίβειας του map matching επειδή εάν υπάρχει χαμηλή ακρίβεια θέσης, η ποιότητα των υπηρεσιών θέσης μειώνεται σημαντικά. Η εργασία [LZZX2009] ασχολείται με το πρόβλημα του σφάλματος δειγματοληψίας επειδή στην πράξη, υφίσταται χαμηλός ρυθμός δειγματοληψίας GPS τροχιών. Στην εργασία αυτή, προτείνεται ένας νέος map-matching αλγόριθμος για χαμηλό ρυθμό δειγματοληψίας σε GPS τροχιές.

Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα με πολλές δυνατότητες, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση του PostGIS. Το PostGIS είναι ένα ανοικτό πρόγραμμα λογισμικού που προσθέτει υποστήριξη για γεωγραφικά αντικείμενα στην αντικείμενο-σχεσιακή βάση δεδομένων PostgreSQL. Επίσης, έχει χρησιμοποιηθεί το pgRouting που είναι επέκταση του PostGIS και της PostgreSQL και δίνει δυνατότητες δρομολόγησης σε γεωγραφικά δεδομένα και άλλες λειτουργίες ανάλυσης του δικτύου. Το pgRouting δίνει την δυνατότητα χωρικών λειτουργιών σε δίκτυα. Για την εύρεση των αποτελεσμάτων έχουν χρησιμοποιηθεί οι συναρτήσεις του pgRouting. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα QGIS.

1.2 Δομή Εργασίας

Το πρώτο κεφάλαιο περιλαμβάνει το εισαγωγικό μέρος της εργασίας δηλαδή, μια σύντομη περιγραφή της εργασίας. Επίσης, την δομή της εργασίας όπου περιγράφουμε περιληπτικά τα κεφάλαια που θα ακολουθήσουν στη συνέχεια της παρούσας εργασίας.

Στο δεύτερο κεφάλαιο παρουσιάζονται οι σχετικές εργασίες. Το κεφάλαιο αυτό, εστιάζει στην χώρο-χρονική μελέτη της συνεχούς κίνησης ανθρώπου (ή γενικότερα κινούμενου αντικείμενου). Αναφέρονται και περιγράφονται έννοιες, των χωρικών και χωροχρονικών δεδομένων, χωροχρονικών βάσεων δεδομένων, κινούμενων αντικείμενων, τροχιάς κίνησης, τροχιές κινούμενων αντικείμενων, η πιθανή περιοχική δραστηριοποίησης κινούμενων αντικείμενων και το πρόβλημα του map matching δηλαδή η προσέγγιση της τροχιάς ενός κινούμενου αντικείμενου πάνω σε ένα χάρτη.

Στο τρίτο κεφαλαίο αναφέρεται η εύρεση της πιθανής περιοχής δραστηριοποίησης (Potential Area of Activity PAA) κινούμενου αντικείμενου. Για την εύρεση της πιθανής περιοχής δραστηριοποίησης κινούμενου αντικείμενου μελετήθηκαν αναλυτικά οι εργασίες [PJ1999], [NVS2010]. Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα με πολλές δυνατότητες, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση PostGIS. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα GIS.

Στο τέταρτο κεφάλαιο παρουσιάζεται το map matching που έχει στόχο να προσεγγίζει όσο πιο πιστά γίνεται την τροχιά ενός κινούμενου αντικείμενου πάνω σε έναν χάρτη. Περιγράφονται οι συναρτήσεις που λύνουν αυτό το πρόβλημα. Χρησιμοποιείται η σχεσιακή βάση δεδομένων ανοικτού κώδικα, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση PostGIS. Επίσης, έχει χρησιμοποιηθεί το pgRouting που είναι επέκταση του PostGIS και της PostgreSQL και δίνει δυνατότητες δρομολόγησης σε γεωγραφικά δεδομένα. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα Quantum GIS.

Στο τέλος ακολουθούν τα συμπεράσματα της παρούσας εργασίας καθώς και μελλοντικές εργασίες και η βιβλιογραφία που χρησιμοποιήθηκε για την άντληση των απαραίτητων πληροφοριών για την ολοκλήρωση της εργασίας.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΚΕΦΑΛΑΙΟ 2

Σχετικές Εργασίες

2.1 Εισαγωγή

Το κεφάλαιο αυτό, εστιάζει στην χώρο-χρονική μελέτη της συνεχούς κίνησης ανθρώπου (ή γενικότερα κινούμενου αντικειμένου). Για την εκπόνηση της παρούσας διπλωματικής εργασίας χρειάστηκε να μελετήσω μια σειρά από επιστημονικά άρθρα πάνω στα κινούμενα αντικείμενα και γενικότερα στα Γεωγραφικά Συστήματα Πληροφοριών (Geographic Information System). Στην συνέχεια θα αναφέρονται και θα περιγράφονται έννοιες, των χωρικών και χωροχρονικών δεδομένων, χωροχρονικών βάσεων δεδομένων, κινούμενων αντικειμένων, τροχιάς κίνησης, τροχιές κινούμενων αντικειμένων, η πιθανή περιοχή δραστηριοποίησης κινούμενων αντικειμένων και τέλος το map-matching πρόβλημα.

Τα χωρικά δεδομένα (spatial data), προσδιορίζουν τα γεωμετρικά χαρακτηριστικά του στοιχείου (θέση, διαστάσεις, σχήμα κ.λπ.) και έχουν άμεση σχέση με τον εντοπισμό του, ενώ τα περιγραφικά δεδομένα ή μη χωρικά δεδομένα (aspatial data) αναφέρονται σε χαρακτηριστικά ή ιδιότητες που αποδίδονται στο συγκεκριμένο στοιχείο του χώρου και δεν σχετίζονται άμεσα με τον εντοπισμό του. Η αναπαράσταση των δεδομένων σε ένα σύστημα γεωγραφικών πληροφοριών γίνεται με δυο τρόπους: διανυσματική ή ψηφιδωτή αναπαράσταση. Τα χωροχρονικά δεδομένα είναι τα δεδομένα που συνδυάζουν χωρική πληροφορία με τη διάσταση του χρόνου. Τα χωροχρονικά δεδομένα είναι μια κατηγορία δεδομένων που περιλαμβάνουν τόσο το χώρο όσο και το χρόνο σαν μεταβλητές.

Οι χωροχρονικές βάσεις δεδομένων έχουν γίνει πολύ σημαντικές τα τελευταία χρόνια, καθώς πολλές εφαρμογές πραγματικού κόσμου, όπως, location based services, Γεωγραφικά συστήματα πληροφοριών (Geographic Information System), κ.λπ. πρέπει να αποθηκεύουν δεδομένα του πραγματικού κόσμου που δείχνουν χωρικά καθώς και χρονικά χαρακτηριστικά. Πολλά αντικείμενα σε πραγματικό κόσμο έχουν χαρακτηριστικά που σχετίζονται με το χώρο και το χρόνο, και η διαχείρισή τους χρησιμοποιώντας τα υπάρχοντα Συστήματα Διαχείρισης Βάσεων Δεδομένων (RDBMS) είναι πολύπλοκη και μη αποδοτική, καθώς αυτά τα αντικείμενα που δείχνουν χωροχρονική συμπεριφορά έχουν πολυδιάστατη φύση. Μια χωροχρονική βάση δεδομένων ενσωματώνει χωρικές, χρονικές και χωροχρονικές έννοιες των βάσεων δεδομένων, και συλλαμβάνει χωρικές και χρονικές πτυχές των δεδομένων. Μια εκτεταμένη έρευνα χωροχρονικών βάσεων δεδομένων γίνεται στην εργασία [AR1999].

Παντού γύρω μας υπάρχουν αντικείμενα που κινούνται. Η κινητικότητα των αντικειμένων δημιουργεί κίνηση. Η κίνηση των αντικειμένων μπορεί να διαχωριστεί σε δύο είδη: κίνηση ελεύθερης τροχιάς και κίνηση περιορισμένης τροχιάς. Χωροχρονικές εφαρμογές που περιλαμβάνουν συνεχώς κινούμενα αντικείμενα είναι όλο και περισσότερο στο επίκεντρο της έρευνας. Θεμελιώδη ζητήματα σε αυτές τις εφαρμογές είναι η απόκτηση και η αναπαράσταση των κινήσεων των αντικειμένων. Μια τροχιά είναι η διαδρομή που ακολουθεί ένα κινούμενο αντικείμενο μέσα στο χώρο σε συνάρτηση με το χρόνο. Μία τροχιά είναι η καταγραφή της κίνησης ενός αντικειμένου, δηλαδή η καταγραφή των θέσεων ενός αντικειμένου σε συγκεκριμένα χρονικά αποτυπώματα. Η κίνηση ενός αντικειμένου αναπαρίστανται στους άξονες x , y (άξονες x , y) και το χρόνο t (t άξονας). Προκειμένου να είναι ακριβής η αναπαράσταση της τροχιάς του, η θέση του αντικειμένου θα πρέπει να καταγράφεται συνεχώς με την πάροδο του χρόνου.

Το πρόβλημα του προσδιορισμού της πιθανής περιοχής δραστηριότητας ενός κινούμενου αντικειμένου βρίσκεται στην προσπάθεια να προσδιορίσει την περιοχή στην οποία το κινούμενο αντικείμενο δυνητικά θα κινηθεί σε μια δεδομένη χρονική περίοδο και με μία δεδομένη μέγιστη ταχύτητα. Δίνοντας ένα αντικείμενο, ένα χρονικό ορόσημο (timestamp) και την μέγιστη ταχύτητα, μπορούμε να υπολογίσουμε μια περιοχή γύρω από το αντικείμενο που θα είναι η περιοχή που ενδεχομένως θα κινηθεί. Αυτό το πρόβλημα μπορεί να παρουσιαστεί με πολλές παραλλαγές.

Η ομάδα Tijds Neutens, Mathias Versichele, Tim Schwanen, κατασκεύασε ένα toolkit το οποίο μελετά και αξιολογεί τις καθημερινές συνήθειες μετακίνησης και γενικά δραστηριοτήτων μέσα στην πόλη, για ένα ή περισσότερα άτομα. Η εφαρμογή αυτή είναι μια αυτόνομη εφαρμογή και μπορεί να στείλει αποτελέσματα στο Google Earth/Maps καθώς, και σε συμβατικό λογισμικό GIS. Μέσω της εφαρμογής αυτής προσπαθούν να συμβάλουν στην ακαδημαϊκή βιβλιογραφία προωθώντας τις προόδους στα συστήματα γεωγραφικών πληροφοριών (GIS). Προσπαθούν να αναπτύξουν μια αναλυτική προσέγγιση για την χώρο-χρονική προσβασιμότητα για κοινές δραστηριότητες, ενώ τα υπάρχοντα μοντέλα της χώρο-χρονικής προσβασιμότητας έχουν την τάση να επικεντρώνονται σε ατομικές δραστηριότητες ή/και να αναλύουν κοινές δραστηριότητες, εάν αυτές είναι ισοδύναμες με τις ατομικές δραστηριότητες. Θέλουν να συμβάλουν όχι μόνο στην

ανάπτυξη ενός συστήματος γεωγραφικών πληροφοριών προσανατολισμένο στον άνθρωπο (Miller, 2007), αλλά και στη διάδοση της ανάλυσης χώρο-χρονικής προσβασιμότητας σε ένα ευρύτερο κοινό. Τους ενδιαφέρει γενικά να απευθυνθούν σε διαφόρων ειδικοτήτων επιστήμονες και να τους διαδώσουν το σκεπτικό των χωροχρονικών εφαρμογών και δυνατοτήτων επεξεργασίας τους. Γι αυτό η εφαρμογή είναι απλή, άμεση και πραγματικού χρόνου. Στην εργασία αυτή δίνουν τις βασικές έννοιες του χωροχρονικού μονοπατιού (space - time path), χωροχρονικού πρίσματος (space - time prism) και της περιοχής δυνητικών μονοπατιών (potential path area) έχοντας υπόψη τη σημασία των περιορισμών για την ολοκλήρωση των ανθρωπίνων δραστηριοτήτων στο χώρο και το χρόνο [NVS2010].

Η παρουσία των κινητών συσκευών έχει δημιουργήσει ένα νέο φάσμα των location-based υπηρεσιών, οι οποίες γίνονται όλο και πιο δημοφιλείς στις μέρες μας. Στο Google Maps, μπορούμε να απολαμβάνουμε την ευκολία των location - based υπηρεσιών, όπως ο σχεδιασμός οδών διαδρομών, η εύρεση εστιατορίων, κλπ. Μελετάμε το πρόβλημα του σχεδιασμού μιας ταξιδιωτικής διαδρομής εξετάζοντας τις ιστορικές τροχιές άλλων ανθρώπων που παράγονται από τις GPS συσκευές. Μια τέτοια συλλογή από τροχιές δίνει συμβουλές για το πώς οι άνθρωποι συνήθως ταξιδεύουν μεταξύ των τοποθεσιών, και στόχος είναι η ανακάλυψη της πιο δημοφιλής διαδρομής από μια συγκεκριμένη τοποθεσία σε μια άλλη. Αυτό είναι εντελώς διαφορετικό από τις υπάρχουσες μεθόδους σχεδιασμού της συντομότερης ή ταχύτερης διαδρομής. Η πιο δημοφιλής διαδρομή είναι ουσιαστικά ένα στατιστικό αποτέλεσμα που προέρχεται από τις πραγματικές ταξιδιωτικές διαδρομές που διεξάγονται από άλλους ανθρώπους στο παρελθόν, και δεν είναι κατ' ανάγκη η συντομότερη διαδρομή.

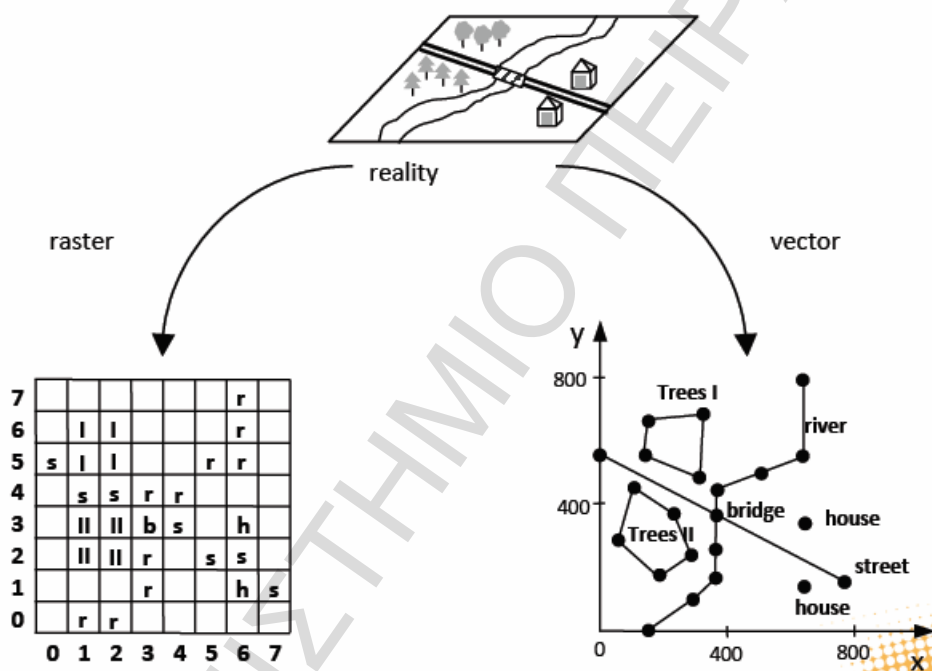
Μια τροχιά είναι μια καταγραφή της διαδρομής από κινούμενα αντικείμενα, όπως είναι τα άτομα, τα οχήματα, τα ζώα και τα φαινόμενα της φύσης. Στην πράξη, ένα μεγάλο μέρος αυτών των τροχιών είναι χαμηλού ρυθμού δειγματοληψίας. Για παράδειγμα, το χρονικό διάστημα μεταξύ διαδοχικών σημείων GPS τροχιών μπορεί να είναι αρκετά λεπτά ή και ώρες. Με ένα τόσο χαμηλό ποσοστό δειγματοληψίας, τα περισσότερα στοιχεία της κίνησης χάνθηκαν, το οποίο καθιστά δύσκολη την αποτελεσματική επεξεργασία. Για παράδειγμα, τα περισσότερα ταξί στις πόλεις είναι εξοπλισμένα με GPS, που τους επιτρέπουν να αναφέρουν χρονοσήμανση θέσεων περιοδικά. Ωστόσο, για την εξοικονόμηση ενέργειας και το κόστος της επικοινωνίας, τα ταξί αναφέρουν τις θέσεις τους σε χαμηλή συχνότητα. Άρα, λεπτομερείς πληροφορίες σχετικά με την ακριβή κίνηση χάνεται και μεγάλη αβεβαιότητα προκύπτει στις διαδρομές. Σε τροχιές, όπως αυτά από τα μεταναστευτικά πουλιά, όπου τα αρχεία τοποθεσιών είναι πολύ αραιά, η αβεβαιότητα γίνεται πιο σημαντική. Αυτό το είδος της αβεβαιότητας θα επηρεάσει την αποτελεσματικότητα και την αποδοτικότητα της μεταγενέστερης διαδικασίας, όπως ευρετηρίαση, αναζήτηση και εξόρυξη. Σκοπός είναι η μείωση της αβεβαιότητας για χαμηλού ρυθμού δειγματοληψίας, δίνεται μια χαμηλού ρυθμού δειγματοληψίας τροχιά, στόχος είναι να προκύψουν πιθανές διαδρομές.

Location based services (LBSs) καλύπτουν ένα ευρύ φάσμα υπηρεσιών παρέχοντας σχετικές πληροφορίες τοποθεσίας. Η πραγματική θέση του χρήστη δεν μπορεί να εμφανίζεται ακριβώς στον ψηφιακό χάρτη. Αυτό είναι ένα από τα σημαντικά θέματα και διαδραματίζει ουσιαστικό ρόλο στις LBSs. Αυτό το φαινόμενο ονομάζεται ως το map-matching πρόβλημα. Το map-matching είναι μια διαδικασία που έχει στόχο να προσεγγίζει όσο πιο πιστά γίνεται την τροχιά ενός κινούμενου αντικείμενου πάνω σε έναν χάρτη. Το πρόβλημα αυτό ονομάζεται map-matching επειδή στόχος είναι η προσπάθεια να «ταιριάξουμε» τη εκτιμωμένη θέση με μια καμπύλη στο χάρτη και να εκτιμηθεί με ακρίβεια η πραγματική θέση του χρήστη [BK1996]. Υπάρχει ένας αριθμός από διαφορετικούς τρόπους για να προσεγγίσει το map matching πρόβλημα, καθένας έχει πλεονεκτήματα και μειονεκτήματα. Έρευνα πάνω στο συγκεκριμένο πρόβλημα έχει γίνει στις εργασίες [WBK2000], [ZG2006], [LZZX2009], [AWB2011].

Στην εργασία [WBK2000] προσεγγίζεται το map matching ως πρόβλημα αναζήτησης και ως στατιστική εκτίμηση. Ο συνδυασμός της απλότητας της απλής προσέγγισης αναζήτησης με ορισμένες από τις ιδέες της στατιστικής προσέγγισης εφαρμόζεται για τέσσερις διαφορετικούς αλγόριθμους. Στην εργασία [AWB2011] μελετώνται τα ζητήματα που αφορούν τις map-matching αβεβαιότητες θέσης και αναλύονται οι παράγοντες που επηρεάζουν το map-matching. Η εργασία [LZZX2009] ασχολείται με το πρόβλημα του σφάλματος δειγματοληψίας επειδή στην πράξη, υφίσταται χαμηλός ρυθμός δειγματοληψίας GPS τροχιών. Στην εργασία αυτή, προτείνεται ένας νέος map-matching αλγόριθμος για χαμηλό ρυθμό δειγματοληψίας σε GPS τροχιές.

2.2 Χωρικά και Χωροχρονικά Δεδομένα

Τα **χωρικά δεδομένα** (spatial data), προσδιορίζουν τα γεωμετρικά χαρακτηριστικά του στοιχείου (θέση, διαστάσεις, σχήμα κ.λπ.) και έχουν άμεση σχέση με τον εντοπισμό του, ενώ τα περιγραφικά δεδομένα ή μη χωρικά δεδομένα (aspatial data) αναφέρονται σε χαρακτηριστικά ή ιδιότητες που αποδίδονται στο συγκεκριμένο στοιχείο του χώρου και δεν σχετίζονται άμεσα με τον εντοπισμό του. Η αναπαράσταση των δεδομένων σε ένα σύστημα γεωγραφικών πληροφοριών γίνεται με δυο τρόπους: διανυσματική (vector) ή ψηφιδωτή αναπαράσταση (raster). Η διανυσματική αναπαριστά τα δεδομένα σύμφωνα με την γεωμετρία τους δηλαδή ως σημεία, γραμμές ή πολύγωνα. Τα σημεία χρησιμοποιούνται για την αναπαράσταση κάποιου σημείου ενδιαφέροντος π.χ. μουσεία, σχολεία, εστιατόρια κ.λπ. Οι γραμμές χρησιμοποιούνται για να αναπαραστήσουν οδούς, ποτάμια κ.λπ. Τα πολύγωνα χρησιμοποιούνται για γεωμετρικές περιοχές π.χ. λίμνες κ.λπ. Η ψηφιδωτή αναπαράσταση είναι ένα πλέγμα που αποτελείται από γραμμές και στήλες. Χαρακτηριστικό παράδειγμα ψηφιδωτών δεδομένων είναι οι δορυφορικές εικόνες.



Εικόνα 2-1: Αναπαράσταση Δεδομένων

Τα **χωροχρονικά δεδομένα** είναι τα δεδομένα που συνδυάζουν χωρική πληροφορία με τη διάσταση του χρόνου. Τα χωροχρονικά δεδομένα είναι μια κατηγορία δεδομένων που περιλαμβάνουν τόσο το χώρο όσο και το χρόνο σαν μεταβλητές. Για το χειρισμό τους ήταν απαραίτητο να αναπτυχθούν νέες μορφές βάσεων δεδομένων έτσι ώστε να είναι δυνατή η επεξεργασία των δύο αυτών μεταβλητών: του χώρου και του χρόνου. Υπάρχουν δυο είδη χωροχρονικών δεδομένων: τα κινούμενα χωροχρονικά δεδομένα που περιλαμβάνουν την έννοια της κίνησης αφού αναφέρονται σε κινούμενα αντικείμενα όπως ένας άνθρωπος ή ένα αυτοκίνητο που κινούνται συνεχώς και τα στατικά χωροχρονικά δεδομένα που δεν περιλαμβάνουν την έννοια της κίνησης καθώς τα αντικείμενα στα οποία αναφέρονται δεν κινούνται όπως ένα κτήριο ή ένα δέντρο ή μια γέφυρα.

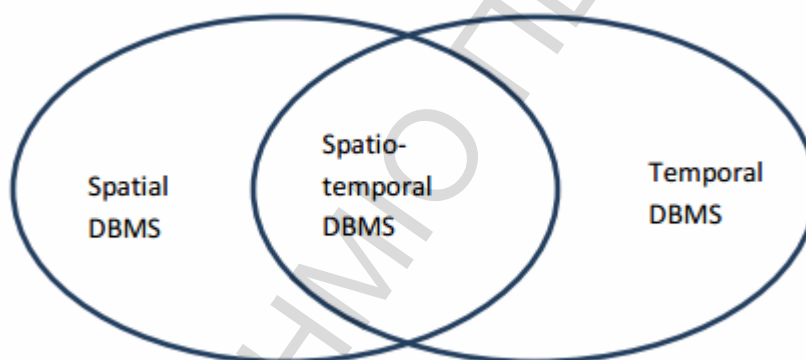
2.3 Χωροχρονικές Βάσεις Δεδομένων

Στα περισσότερα συστήματα βάσεων δεδομένων μοντελοποιείται μια συγκεκριμένη κατάσταση του πραγματικού κόσμου, για παράδειγμα η βάση δεδομένων των φοιτητών ενός πανεπιστημίου. Σε κάποιες περιπτώσεις όμως είναι απαραίτητο να αποθηκεύονται και να

ανακτούνται πληροφορίες σχετικά με καταστάσεις του πραγματικού κόσμου κατά το πέρασμα του χρόνου. Οι βάσεις δεδομένων που μοντελοποιούν συγκεκριμένες καταστάσεις του πραγματικού κόσμου κατά το πέρασμα του χρόνου ονομάζονται χωροχρονικές βάσεις δεδομένων [URL1].

Οι χωροχρονικές βάσεις δεδομένων έχουν γίνει πολύ σημαντικές τα τελευταία χρόνια, καθώς πολλές εφαρμογές πραγματικού κόσμου, όπως, location based services, Γεωγραφικά συστήματα πληροφοριών (Geographic Information System), κ.λπ. πρέπει να αποθηκεύουν δεδομένα του πραγματικού κόσμου που δείχνουν χωρικά καθώς και χρονικά χαρακτηριστικά. Πολλά αντικείμενα σε πραγματικό κόσμο έχουν χαρακτηριστικά που σχετίζονται με το χώρο και το χρόνο, και η διαχείρισή τους χρησιμοποιώντας τα υπάρχοντα Συστήματα Διαχείρισης Βάσεων Δεδομένων (RDBMS) είναι πολύπλοκη και μη αποδοτική, καθώς αυτά τα αντικείμενα που δείχνουν χωροχρονική συμπεριφορά έχουν πολυδιάστατη φύση.

Μια χωροχρονική βάση δεδομένων είναι μια βάση δεδομένων που διαχειρίζεται χωρικές και χρονικές πληροφορίες. Οι χωροχρονικές βάσεις δεδομένων είναι μια επέκταση των χωρικών βάσεων δεδομένων. Μια χωροχρονική βάση δεδομένων ενσωματώνει χωρικές, χρονικές και χωροχρονικές έννοιες των βάσεων δεδομένων, και συλλαμβάνει χωρικές και χρονικές πτυχές των δεδομένων. Μια εκτεταμένη έρευνα χωροχρονικών βάσεων δεδομένων γίνεται στην εργασία [AR1999].



Εικόνα 2-2: Σχέση μεταξύ χωρικών, χρονικών και χωροχρονικών DBMS

Κοινά παραδείγματα [URL2]:

- Παρακολούθηση των κινούμενων αντικειμένων, το οποίο τυπικά μπορεί να καταλαμβάνει μόνο μια ενιαία θέση σε μια δεδομένη στιγμή.
- Μια βάση δεδομένων των ασύρματων δικτύων επικοινωνίας, που μπορεί να υπάρξει μόνο για ένα σύντομο χρονικό διάστημα σε μια γεωγραφική περιοχή.
- Ένας δείκτης των ειδών σε μια δεδομένη γεωγραφική περιοχή, όπου με το πέρασμα του χρόνου επιπλέον είδη μπορούν να εισαχθούν ή υφιστάμενα είδη να μεταναστεύσουν ή να πεθάνουν.
- Ιστορική παρακολούθηση της δραστηριότητας των τεκτονικών πλακών.

Τα τελευταία χρόνια ένα πολύ μεγάλο αριθμό εφαρμογών έχει ενσωματώσει χωροχρονικές έννοιες όπως [URL1]:

- Location based services έχουν στόχο τα κινούμενα πρόσωπα για να τους παρέχουν στοιχεία, σύμφωνα με τα ενδιαφέροντα και τη θέση τους.
- Στην περίπτωση της πρόληψης μπουτιλιάρισματος, το Global Positioning System (GPS) θα μπορούσε να προτείνει νέες τροχιές με βάση τον όγκο της κυκλοφορίας και της συμφόρησης στις γραμμές κατά μήκος του δρόμου ενός ατόμου.

- Η έννοια αυτή χρησιμοποιείται ευρέως στα Γεωγραφικά Πληροφορικά Συστήματα (GIS), συστημάτων περιβαλλοντικής πληροφόρησης και πολυμέσων.
- Επίσης εφαρμογές όπως η παγκόσμια κλιματική αλλαγή, οι δημογραφικές πληροφορίες, τα αρχεία της υγείας.

2.4 Κινούμενα Αντικείμενα

Παντού γύρω μας υπάρχουν αντικείμενα τα οποία κινούνται. Η κινητικότητα των αντικειμένων δημιουργεί κίνηση. Η κίνηση των αντικειμένων μπορεί να διαχωριστεί σε δύο είδη: κίνηση ελεύθερης τροχιάς και κίνηση περιορισμένης τροχιάς. Ένα κινούμενο αντικείμενο έχει ελεύθερη τροχιά όταν υπάρχουν λίγοι περιορισμοί της κίνησής του στον κ-διάστατο χώρο, π.χ. ένας άνθρωπος που κινείται στη πόλη. Αντίθετα ένα κινούμενο αντικείμενο έχει περιορισμένη τροχιά όταν υπάρχουν αυστηροί περιορισμοί στην κίνησή του, όπως π.χ. η κίνηση ενός τρένου στις ράγες. Ενδεχομένως δε με βάση την κίνηση που ήδη έχει κάνει το αντικείμενο στο παρελθόν, να μπορεί να γίνει και πρόβλεψη σχετικά με κάποιες από τις μελλοντικές του κινήσεις.

Χωροχρονικές εφαρμογές που περιλαμβάνουν συνεχώς κινούμενα αντικείμενα είναι όλο και περισσότερες στο επίκεντρο της έρευνας. Έρευνα σε κινούμενα αντικείμενα έχει δημοσιευθεί στις εργασίες [GBEJ2000], [PJ1999]. Στις εφαρμογές αυτές, το μόνο που είναι διαθέσιμο είναι η θέση των αντικειμένων σε διάφορα χρονικά διαστήματα. Η αναπαράσταση των συνεχώς μεταβαλλόμενων θέσεων των αντικειμένων είναι σημαντική σε αυτές τις εφαρμογές. Θεμελιώδη ζητήματα σε αυτές τις εφαρμογές είναι η απόκτηση και η αναπαράσταση των κινήσεων των αντικειμένων.

Το Global Positioning System (GPS), είναι η τεχνολογία που χρησιμοποιείται για τη δειγματοληψία των θέσεων των κινούμενων αντικειμένων. Ο δειγματισμός των θέσεων, θα γίνεται με GPS και θα εφαρμόζεται η παρεμβολή για να υπολογίζονται οι ενδιάμεσες θέσεις. Με το GPS μπορούμε να παίρνουμε πληροφορίες για τη θέση αντικειμένων (από τέσσερις ή και περισσότερους δορυφόρους, όπου λύνονται 4 εξισώσεις που εμπεριέχουν μεγέθη όπως γεωγραφικό πλάτος, γεωγραφικό μήκος, ύψος και χρόνος), αλλά και τις σχετικές τους αποστάσεις από αντικείμενα ενδιαφέροντος αλλά και ομάδων ομοειδών αντικειμένων.

Ως αποτέλεσμα, η καταγραφή των κινήσεων των αντικειμένων όπως αποθηκεύονται στη βάση δεδομένων διαφέρει από την πραγματική κίνηση. Όσο μεγαλώνει ο ρυθμός δειγματισμού, τόσο οι θέσεις δειγματισμού ανταποκρίνονται καλύτερα στην πραγματική κίνηση του αντικειμένου και έτσι το λάθος που οφείλεται στον δειγματισμό, ελαττώνεται.

Έτσι, τα συστήματα χωροχρονικών βάσεων δεδομένων αποτελούν στην ουσία συστήματα κινούμενων αντικειμένων (moving objects) με ιδιαίτερο και ερευνητικά γόνιμο ενδιαφέρον, προσφέροντας τη δυνατότητα να συγκεραστούν ιδέες από τα συγγενικά πεδία των χωρικών και των χρονικών βάσεων δεδομένων. Ως χαρακτηριστικό παράδειγμα θα μπορούσε να χρησιμεύσει ένα υποθετικό σύστημα διευκόλυνσης της κυκλοφορίας των οχημάτων στο οδικό δίκτυο μιας πυκνοκατοικημένης περιοχής, και ειδικότερα η διαχείριση στόλου οχημάτων (fleet management) για μεγάλους προμηθευτές αγαθών ή υπηρεσιών. Τα οχήματα φέρουν συσκευές GPS και μεταδίδουν ανά τακτά χρονικά διαστήματα (μερικά δευτερόλεπτα) τη γεωγραφική τους θέση σε ένα κεντρικό υπολογιστικό σύστημα, μέσω ειδικών ραδιοφωνικών συχνοτήτων ή σημάτων κινητής τηλεφωνίας (Global System for Mobile communications, GSM). Στο κέντρο ελέγχου, γίνεται η επεξεργασία των δεδομένων και η αξιοποίησή τους ανάλογα με τους σκοπούς του συστήματος (για εξοικονόμηση πόρων, κάλυψη απρόοπτων αναγκών, εποπτεία τήρησης δρομολογίων κλπ.) [URL3].

Η κίνηση των αντικειμένων είναι πολύ πιθανόν να διέπεται από περιορισμούς, αφού δεν είναι λογικό να υποτεθεί ότι αυτά κινούνται ανεμπόδια σε έναν δισδιάστατο ευκλείδειο χώρο. Έτσι, οι κινήσεις μπορούν να διακριθούν σε τρεις κύριες κατηγορίες:

- Κίνηση χωρίς περιορισμούς: εξετάζεται κυρίως από τις μεθόδους προσπέλασης χωροχρονικών στοιχείων και προϋποθέτει ως υπόβαθρο ένα τέλειο ευκλείδειο επίπεδο, για παράδειγμα η κίνηση πλοίων στη θάλασσα.
- Κίνηση υπό περιορισμούς: (όπως η κίνηση των πεζών ή των οχημάτων) θεωρείτε ότι υφίστανται φυσικά εμπόδια που δυσχεραίνουν ή απαγορεύουν τη μετάβαση (κτίρια ή

υδάτινες επιφάνειες). Αυτά τα εμπόδια θεωρούνται ως νεκρός χώρος, επειδή υπάρχουν περιορισμένες ή ανύπαρκτες δυνατότητες κίνησης.

- Κίνηση σε δίκτυα μεταφορών: η μετατόπιση των αντικειμένων συμβαίνει αποκλειστικά πάνω στο δίκτυο και όχι γενικά στο δισδιάστατο επίπεδο, για παράδειγμα το οδικό δίκτυο. Οι χρήστες δεν ενδιαφέρονται τόσο πολύ για τις απόλυτες συντεταγμένες της θέσης των αντικειμένων, σημασία έχει κυρίως η σχετική θέση των κινούμενων αντικειμένων ως προς τα τμήματα του δικτύου [URL3].

2.5 Τροχιά Κινούμενων Αντικειμένων

Μια τροχιά (trajectory) είναι η διαδρομή που ακολουθεί ένα κινούμενο αντικείμενο μέσα στο χώρο σε συνάρτηση με το χρόνο. Για παράδειγμα, το αντικείμενο μπορεί να είναι ένας δορυφόρος ή μια σφαίρα ή ένας άνθρωπος ή ένα αυτοκίνητο κλπ. Μια τροχιά μπορεί να περιγραφεί μαθηματικά είτε με τη γεωμετρία της διαδρομής, ή τη θέση του αντικειμένου στην πάροδο του χρόνου.

Μία τροχιά μπορεί να οριστεί ως μία συνάρτηση από το χρονικό $I \subseteq \mathbb{R}$ πεδίο στο γεωγραφικό χώρο \mathbb{R}^2 . Πιο συγκεκριμένα, μία τροχιά T είναι μία συνεχής αναπαράσταση από το χρονικό $I \subseteq \mathbb{R}$ στο χωρικό πεδίο (\mathbb{R}^2 στο δισδιάστατο επίπεδο):

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

και

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

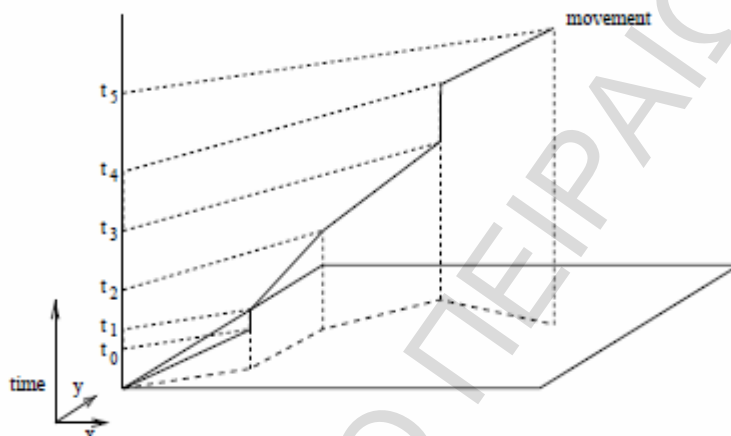
Μία τροχιά είναι η καταγραφή της κίνησης ενός αντικειμένου, δηλαδή η καταγραφή των θέσεων ενός αντικειμένου σε συγκεκριμένα χρονικά αποτυπώματα, ενώ η πραγματική τροχιά αποτελείται στη γενική περίπτωση από μία καμπύλη, η οποία, αποτελείται από μία σειρά από δειγματοληπτούμενα σημεία. Έτσι, οι τροχιές των κινούμενων σημείων συχνά ορίζονται ως ακολουθίες τριάδων (x, y, t) :

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

όπου, $x_i, t_i, y_i \in \mathbb{R}$, και $t_1 < t_2 \dots < t_n$, και η καμπύλη της τροχιάς δημιουργείται κατά προσέγγιση εφαρμόζοντας χωροχρονικές μεθόδους παρεμβολής στο σύνολο των τυχαίων σημείων [URL4].

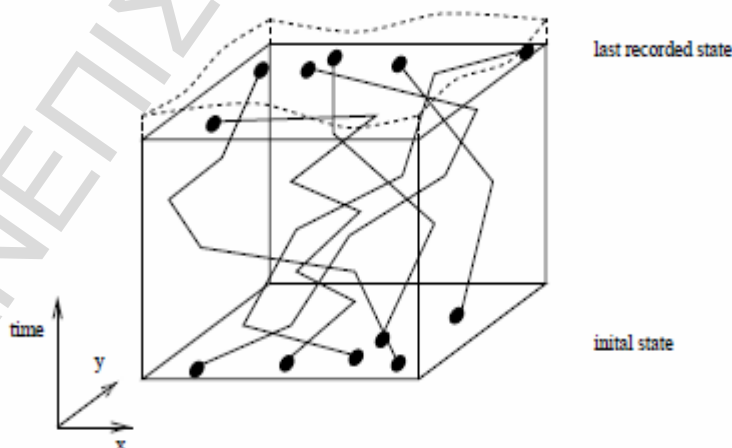
Προκειμένου να καταγράψουμε την κίνηση ενός αντικειμένου, θα έπρεπε να γνωρίζουμε τη θέση ανά πάσα στιγμή, δηλαδή, σε συνεχή βάση. Ωστόσο, το GPS μας επιτρέπει να δειγματίζουμε τη θέση ενός αντικειμένου, δηλαδή, να ληφθεί η θέση σε διακριτές χρονικές στιγμές. Η κίνηση ενός αντικειμένου αναπαρίστανται στους άξονες x και y (άξονες x, y) και το χρόνο t (t άξονα). Καταγράφοντας τις διαδοχικές θέσεις ενός κινούμενου αντικειμένου προκύπτει η τροχιά του. Προκειμένου να είναι ακριβής η αναπαράσταση της τροχιάς του, η θέση του αντικειμένου θα πρέπει να καταγράφεται συνεχώς με την πάροδο του χρόνου. Κάτι τέτοιο είναι πρακτικά ανέφικτο, αφού οι υπάρχουσες τεχνολογικές δυνατότητες (τηλεπικοινωνίες, GPS) παρέχουν μόνο ένα δείγμα των θέσεων του αντικειμένου. Έτσι, πρέπει οι επιδιώξεις της μοντελοποίησης να συμβιβαστούν με το γεγονός ότι οι σημειακές θέσεις των αντικειμένων λαμβάνονται από τις συσκευές καταγραφής σε διακριτές χρονικές στιγμές, επομένως η τροχιά πρέπει να αντιμετωπίζεται ως δειγματοληψία της πραγματικής κίνησης. Οι συντεταγμένες των θέσεων πρέπει να είναι άρρηκτα συνδεδεμένες με τα χρονικά ορόσημα (timestamps) όπου σημειώθηκαν, οπότε κάθε σημείο της τροχιάς προσδιορίζεται στον τρισδιάστατο χώρο (x, y, t) [URL3].

Συνήθως γίνεται απόπειρα αναπαράστασης ολόκληρης της κίνησης των αντικειμένων με κάποια τεχνική όπως είναι η μέθοδος της παρεμβολής (interpolation). Όσον αφορά στις περισσότερες σχετικές εφαρμογές, η μέθοδος της γραμμικής παρεμβολής εμφανίζεται επαρκής. Κατά συνέπεια, τα γνωστά σημεία θεωρούνται ως άκρα διαδοχικών ευθυγράμμων τμημάτων, σχηματίζοντας μια τεθλασμένη πολυγραμμή (polyline) που απεικονίζει την τροχιά του αντικειμένου στον τρισδιάστατο χώρο (x, y, t) . Αυτή η πολυγραμμή θα αποθηκευτεί τελικά σε μια (χωρική) βάση δεδομένων. Οι συντεταγμένες του χώρου και του χρόνου συνδυάζονται ώστε να αποτελέσουν κοινό σύστημα αναφοράς. Η διακεκομμένη γραμμή συμβολίζει την προβολή (ίχνος) της τροχιάς στο χωρικό επίπεδο (x, y) [PJ1999].



Εικόνα 2-3: Τροχιά ενός κινούμενου αντικείμενου

Η διαδρομή ενός κινούμενου αντικείμενου ονομάζεται τροχιά, εναλλακτικά κίνηση. Μια πρώτη προσέγγιση για να αντιπροσωπευθούν οι κινήσεις των αντικειμένων θα είναι να αποθηκεύονται τα δείγματα των θέσεων τους. Η απλούστερη προσέγγιση είναι η χρήση γραμμικής παρεμβολής. Οι θέσεις δειγματοληψίας γίνονται τα τελικά σημεία της γραμμής και η κίνηση ενός αντικείμενου αντιπροσωπεύεται από ένα polyline σε τρισδιάστατο χώρο όπως φαίνεται στην παρακάτω εικόνα [PJ1999].



Εικόνα 2-4: Τροχιές στο χωροχρονικό διάστημα

Η παραπάνω εικόνα δείχνει το χωροχρονικό διάστημα και αρκετές τροχιές. Ο χρόνος κινείται στην κατεύθυνση προς τα πάνω, και στην κορυφή του κύβου είναι ο χρόνος του πιο πρόσφατου δείγματος θέσης.

Άρα, πρέπει να θεωρήσουμε τις διακριτές θέσεις που παίρνει το αντικείμενο σε συνάρτηση με το χρόνο. Όλη η διαδρομή (trajectory) του αντικειμένου μπορεί να θεωρηθεί σαν σύνθεση πολλών μικρών διαδρομών (segments), που με γραμμική παρεμβολή μπορούμε να τις εμπλουτίσουμε με όλες τις ενδιάμεσες θέσεις που παίρνει το αντικείμενο. Οι διαδρομές αυτές είναι ανακριβείς και αυτό οφείλεται τόσο στην διαδικασία της μέτρησης που δειγματίζει τις θέσεις του αντικειμένου, όσο και της ίδιας της προσέγγισης δειγματισμού που γίνεται. Πάντως εμείς θα θεωρήσουμε ότι οι χρονικές στιγμές που λαμβάνονται οι θέσεις του αντικειμένου, είναι ακριβείς. Ο δειγματισμός των θέσεων, θα γίνεται με GPS. Με το GPS μπορούμε να παίρνουμε πληροφορίες για τη θέση αντικειμένων αλλά και τις σχετικές τους αποστάσεις από αντικείμενα ενδιαφέροντος αλλά και ομάδων ομοειδών αντικειμένων.

Σε μία εφαρμογή, εξετάσουμε μόνο τα αντικείμενα με παρόμοια χαρακτηριστικά κίνησης, όπως η ταχύτητα και το εύρος. Αντικείμενα με διαφορετικά χαρακτηριστικά περιλαμβάνουν ανθρώπους, αυτοκίνητα, αεροπλάνα. Επίσης, για περισσότερα από ένα αντικείμενα, θεωρούμε ότι έχουν παρόμοια χαρακτηριστικά στην ταχύτητα και το εύρος που κινούνται. Π.χ. αν υποθέσουμε ότι παρακολουθούμε την κίνηση, δύο αντικειμένων, αυτά δεν θα μπορούσαν να είναι ένας άνθρωπος και ένα αεροπλάνο [PJ1999].

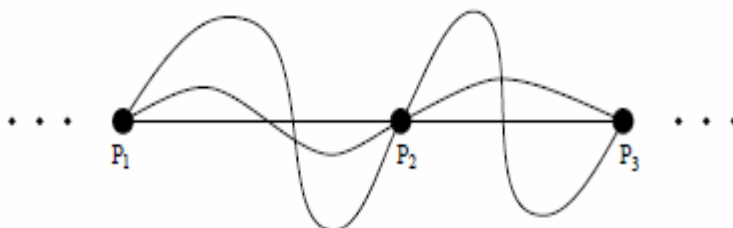
2.6 Πιθανή Περιοχή Δραστηριοποίησης Κινούμενων Αντικειμένων

Το πρόβλημα του προσδιορισμού της πιθανής περιοχής δραστηριότητας ενός κινούμενου αντικειμένου βρίσκεται στην προσπάθεια να προσδιορίσει την περιοχή στην οποία το κινούμενο αντικείμενο δυνητικά θα κινηθεί σε μια δεδομένη χρονική περίοδο και με μία δεδομένη μέγιστη ταχύτητα. Δίνοντας ένα αντικείμενο, ένα χρονικό ορόσημο (timestamp) και την μέγιστη ταχύτητα του αντικειμένου, μπορούμε να υπολογίσουμε μια περιοχή γύρω από το αντικείμενο που θα είναι η περιοχή που ενδεχομένως θα κινηθεί.

Αυτό το πρόβλημα μπορεί να παρουσιαστεί με πολλές παραλλαγές. Η γενική προσέγγιση αυτού του προβλήματος είναι να λαμβάνει υπ' όψιν: ένα κινούμενο αντικείμενο O , n timestamps t_i , $i = 1 \dots m \dots n$ (n διαδοχικά δείγματα), n χωρικές θέσεις x_i, y_i , $i = 1 \dots m \dots n$ (οι θέσεις του αντικειμένου στο t_i) και η παρεμβολή των χωρικών θέσεων (η οποία είναι η πραγματική κίνηση του αντικειμένου), δύο timestamps t_i και t_k και η μέγιστη ταχύτητα M που μπορεί να επιτευχθεί από το συγκεκριμένο κινούμενο αντικείμενο. Η λύση του προβλήματος είναι να προσπαθήσουμε να βρούμε την πιθανή περιοχή της κίνησης, θεωρώντας ότι το κινούμενο αντικείμενο έχει μέγιστη ταχύτητα και ότι θα ξεκινήσει από το χωροχρονικό σημείο x_m, y_m, t_m και να τελειώσει στο χωροχρονικό σημείο $x_{m+1}, y_{m+1}, t_{m+1}$.

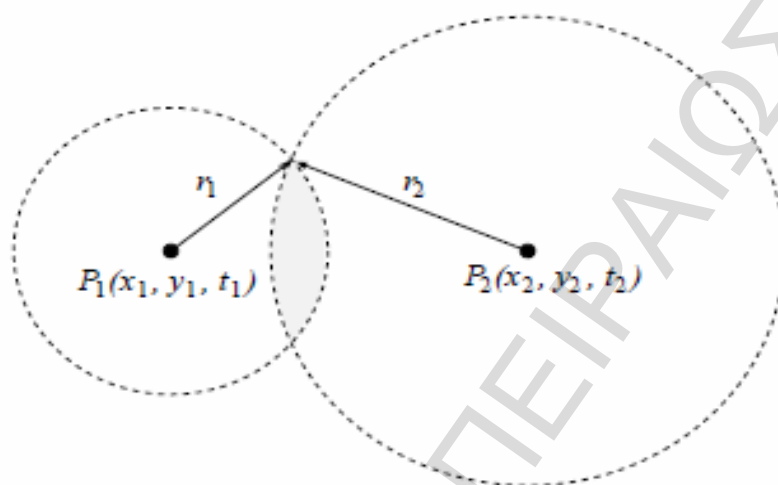
Η δεύτερη προσέγγιση αυτού του προβλήματος είναι να προσπαθήσουμε να βρούμε την πιθανή περιοχή της κίνησης του κινούμενου αντικειμένου στο χωροχρονικό σημείο t_j , θεωρώντας ότι έχει μέγιστη ταχύτητα M και ότι θα ξεκινήσει από το χωροχρονικό σημείο x_m, y_m, t_m και πρέπει να έχει ολοκληρωθεί στο χωροχρονικό σημείο $x_{m+1}, y_{m+1}, t_{m+1}$. Στην πραγματικότητα, το πρόβλημα που υπάρχει προσπαθεί να προσδιορίσει την πιθανή περιοχή της κίνησης από t_m προς t_j και από t_j προς t_{m+1} έχοντας μέγιστη ταχύτητα M . Το σημείο τομής των δύο κύκλων είναι η δυνητική περιοχή δραστηριότητας για το κινούμενο αντικείμενο στο t_j . Παρακάτω θα δούμε αναλυτικά τον υπολογισμό της περιοχής αυτής.

Εξετάζουμε με ποιο τρόπο τα δείγματα θέσης μοιάζουν με την αληθινή κίνηση του αντικειμένου. Στην Εικόνα 2-5 εμφανίζονται τρεις τροχιές. Κάθε μια δίνεται από τις τρεις θέσεις μέτρησης P_1 έως P_3 [PJ1999].



Εικόνα 2-5: Πιθανές τροχιές κινούμενου αντικειμένου

Λαμβάνοντας υπόψη την τροχιά σε ένα χρονικό διάστημα $[t_1, t_2]$, που οριοθετείται από διαδοχικά δείγματα, γνωρίζουμε δύο θέσεις, P_1 και P_2 , καθώς και τη μέγιστη ταχύτητα του αντικειμένου, v_m , τότε για κάποια ενδιάμεση χρονική στιγμή t_x στο χρονικό διάστημα κίνησης του, το αντικείμενο θα κινηθεί στην τομή των κύκλων που ορίζονται με ακτίνες $r_1=(t_1+t_x)*v_m$ γύρω από το P_1 (ο μικρότερος κύκλος) και $r_2=(t_2-t_x)*v_m$ γύρω από το P_2 (ο μεγαλύτερος κύκλος). Η τομή αυτή φαίνεται στην παρακάτω εικόνα [PJ1999].



Εικόνα 2-6: Η τομή των δύο κύκλων

Έτσι, κατά το χρόνο t_x για μία συγκεκριμένη θέση μπορεί να περιγραφεί με τη συνάρτηση πιθανότητας ($P_2(x,y)$) που φαίνεται παρακάτω, όπου r_1 και r_2 είναι οι δύο ακτίνες που περιγράφονται παραπάνω, το s είναι η απόσταση μεταξύ των P_1 και P_2 , και A υποδηλώνει την περιοχή της τομής των δύο κύκλων.

$$P_2(x,y) = \begin{cases} \frac{1}{A} & \text{for } x^2 + y^2 \leq r_1^2 \wedge (x-s)^2 + y^2 \leq r_2^2 \\ 0 & \text{otherwise} \end{cases}$$

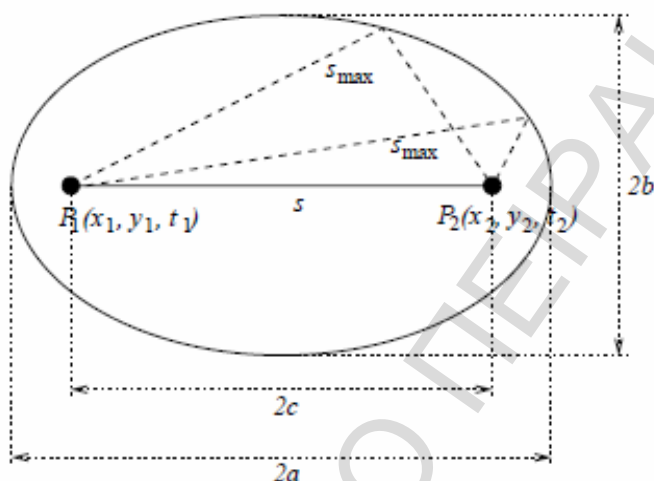
Αντικαθιστώντας $v_m(t_1 + t_x)$ και $v_m(t_2 - t_x)$ τις ακτίνες R_1 και R_2 , αντίστοιχα, η συνάρτηση πιθανότητας φαίνεται στην παρακάτω εξίσωση. Οι παράμετροι περιγράφονται στον Πίνακα 1.

$$P_2(x,y) = \begin{cases} \frac{1}{A} & \text{for } x^2 + y^2 \leq (v_m(t_1 + t_x))^2 \wedge \\ & (x-s)^2 + y^2 \leq (v_m(t_2 - t_x))^2 \\ 0 & \text{otherwise} \end{cases}$$

| | |
|-------|--|
| v_m | Μέγιστη ταχύτητα του κινούμενου αντικειμένου |
| t_x | Ο χρόνος για μία συγκεκριμένη θέση |
| t_1 | Ο χρόνος μέτρησης της πρώτης θέσης |
| t_2 | Ο χρόνος μέτρησης της δεύτερης θέσης |
| s | Η απόσταση ανάμεσα στα δυο σημεία |
| A | Η περιοχή της τομής των δύο κύκλων |

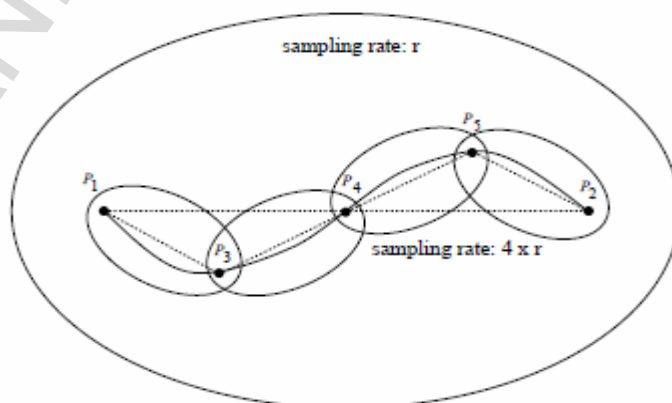
Πίνακας 2-1: Παράμετροι της συνάρτησης πιθανότητας P_2

Παρατηρούμε ότι τα σημεία τομής των δύο κύκλων, βρίσκονται σε μία έλλειψη με τις θέσεις P1 και P2 ως σταθερά σημεία της, όπως φαίνεται στην Εικόνα 2-7. Το μήκος του μεγάλου ημιάξονα είναι $2a = r_1 + r_2$. Μια έλλειψη είναι μία καμπύλη που αποτελείται από όλα τα σημεία στο επίπεδο των οποίων το άθροισμα των αποστάσεων, r_1 και r_2 , από δύο σταθερά σημεία, τα P1 και P2 διαχωρίζονται από μία απόσταση $2c$. Το μέτρο $2c$ μπορεί να ορίζεται ως η απόσταση μεταξύ P1 και P2, ενώ $2a$ είναι ο μεγάλος άξονας της έλλειψης. Το πάχος της έλλειψης (ο μικρός άξονας), $2b$, προσδιορίζεται από την εξίσωση $b^2 = a^2 - c^2$. Αυτό σημαίνει ότι όσο μικρότερη είναι η διαφορά μεταξύ της παρατηρούμενης απόστασης, $2c$, και της μέγιστης απόστασης, $2a$, τόσο λεπτότερη η έλλειψη.



Εικόνα 2-7: Έλλειψη

Ρυθμός δειγματοληψίας. Όσο μεγαλύτερη είναι η έλλειψη, τόσο πιο μεγάλο το λάθος μέτρησης. Για να μικρύνει η έλλειψη, πρέπει να έρθουν πιο κοντά τα διάφορα σημεία δειγματοληψίας και να μεγαλώσει ο ρυθμός δειγματοληψίας. Για παράδειγμα στην Εικόνα 2-8 δείχνουμε την πραγματική τροχιά ενός κινούμενου αντικειμένου. Ως πρώτο βήμα, δειγματίζουμε την κίνηση του αντικειμένου στη θέση P1 και P2. Ο χρόνος μεταξύ των δειγμάτων είναι 10 δευτερόλεπτα. Η μικρότερη απόσταση από P1 σε P2 είναι 300 μέτρα. Έτσι, το αντικείμενο κινείται με μία ταχύτητα u των 30m/s. Γνωρίζουμε τη μέγιστη ταχύτητα του αντικειμένου που είναι 42m/s, σχεδιάζουμε μια έλλειψη γύρω από τη γραμμή που προσεγγίζει την κίνηση. Η έλλειψη έχει εκκεντρότητα $2c=300$, ο μεγάλος άξονας $2a=U_{max} * \Delta t = 42m/s * 10s = 420m$, και ένα μικρό άξονα $2b=294m$. Τετραπλασιάζοντας το ρυθμό δειγματοληψίας, οδηγεί σε έλλειψη που έχει εκκεντρότητα $2c=80m$, ένα μεγάλο άξονα $2a=105m$, και ένα μικρό άξονα $2b=68m$. Άρα, όσο μεγαλώνει ο ρυθμός δειγματοληψίας, τόσο οι θέσεις δειγματοληψίας ανταποκρίνονται καλύτερα στην πραγματική κίνηση του αντικειμένου [PJ1999].



Εικόνα 2-8: Ρυθμός δειγματοληψίας

Μια άλλη προσέγγιση αυτού του προβλήματος είναι να προσπαθήσουμε να βρούμε την πιθανή περιοχή της κίνησης του κινούμενου αντικείμενου ανάμεσα στο χωροχρονικό σημείο t_j και t_k , θεωρώντας ότι έχει μέγιστη ταχύτητα M και ότι θα ξεκινήσει από το χωροχρονικό σημείο $x_m y_m t_m$ και πρέπει να έχει ολοκληρωθεί στο χωροχρονικό σημείο $x_{m+2} y_{m+2} t_{m+2}$. Στην πραγματικότητα, το πρόβλημα που υπάρχει προσπαθεί να προσδιορίσει την πιθανή περιοχή της κίνησης από t_j προς t_k έχοντας μέγιστη ταχύτητα M . Το σημείο τομής των δύο ελλείψεων είναι η δυνητική περιοχή δραστηριότητας για το κινούμενο αντικείμενο O από t_j προς t_k . Όπως, φαίνεται και στην Εικόνα 2-8 τα σημεία τομής των μικρών ελλείψεων είναι μια πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικείμενου. Τα παράδειγμα της εικόνας 2-8 θα υλοποιηθεί στο τρίτο κεφάλαιο της εργασίας.

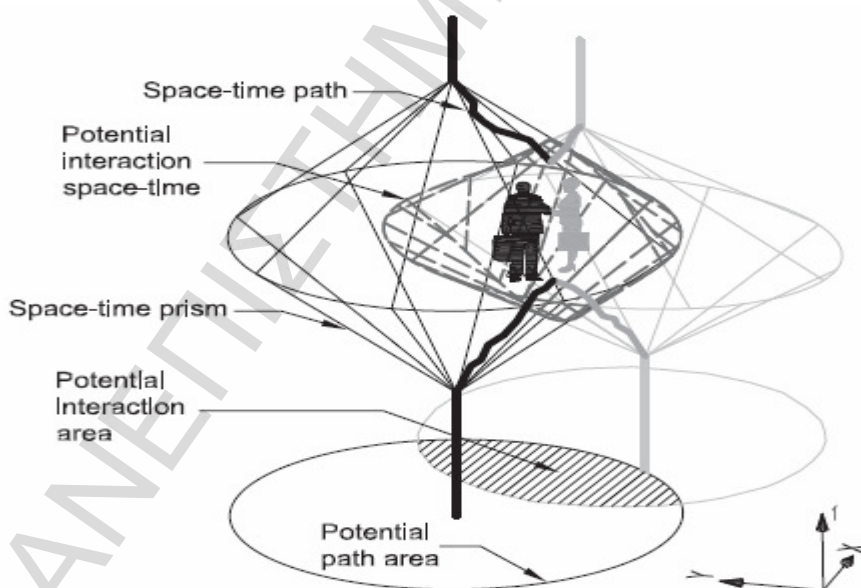
Τέλος, μια άλλη προσέγγιση αυτού του προβλήματος είναι να προσπαθήσουμε να βρούμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικείμενου σε ένα δίκτυο. Η απλή προσέγγιση αυτού του προβλήματος είναι να λαμβάνει υπ' όψιν: Ένα οδικό δίκτυο, ένα κινούμενο αντικείμενο O , n timestamps $t_i, i = 1 \dots m \dots n$ (n διαδοχικά δείγματα), n χωρικές θέσεις $x_i y_i, i = 1 \dots m \dots n$ (οι θέσεις του αντικείμενου στο t_i) και η παρεμβολή των χωρικών θέσεων (η οποία είναι η πραγματική κίνηση του αντικείμενου), δύο timestamps t_j και t_k και η μέγιστη ταχύτητα M που μπορεί να επιτευχθεί από το συγκεκριμένο κινούμενο αντικείμενο. Η λύση του προβλήματος (που σχηματίζεται με τον παραπάνω τρόπο) είναι να προσπαθήσουμε να βρούμε τα πιθανά μονοπάτια της κίνησης, θεωρώντας ότι το κινούμενο αντικείμενο έχει μέγιστη ταχύτητα και ότι θα ξεκινήσει από το χωροχρονικό σημείο $x_m y_m t_m$ και να τελειώσει στο χωροχρονικό σημείο $x_{m+1} y_{m+1} t_{m+1}$. Μπορούμε να παρατηρήσουμε ότι υπάρχει ένα άνω όριο για το δίκτυο περιορισμένης περιοχής δραστηριότητας που είναι η Ευκλείδεια πιθανή περιοχή δραστηριοποίησης. Σε αυτή την περίπτωση είναι η έλλειψη που έχουμε ήδη δει. Μια αφελής προσέγγιση του προβλήματος είναι να προσπαθήσουμε να βρούμε όλες τις πιθανές διαδρομές από $x_m y_m t_m$ προς $x_{m+1} y_{m+1} t_{m+1}$ που δεν είναι μεγαλύτερο από το $MaxDist$ η οποία είναι η απόσταση που διανύεται από το κινούμενο αντικείμενο όταν ταξιδεύει με τη μέγιστη ταχύτητα. Λαμβάνοντας υπόψη την παρατήρηση σχετικά με το ανώτατο όριο του δικτύου περιορισμένης περιοχής δραστηριότητας, μπορούμε να κλαδέψετε το δίκτυο (η οποία είναι στην πραγματικότητα ένα γράφημα) και να κρατήσει μόνο τις ακμές και κορυφές που βρισκόταν στην Ευκλείδεια πιθανή περιοχή δραστηριοποίησης, η οποία στην παρούσα υπόθεση είναι μια έλλειψη.

2.7 Σημεία Ενδιαφέροντος σε Περιοχή Δραστηριοποίησης εν Κινήσει Χρηστών

Η ομάδα Tijds Neutens, Mathias Versichele, Tim Schwanen, κατασκεύασε ένα toolkit το οποίο μελετά και αξιολογεί τις καθημερινές συνήθειες μετακίνησης και γενικά δραστηριοτήτων μέσα στην πόλη, για ένα ή και για περισσότερα άτομα. Η εφαρμογή αυτή είναι μια αυτόνομη εφαρμογή και μπορεί να στείλει αποτελέσματα στο Google Earth/Maps καθώς, και σε συμβατικό λογισμικό GIS. Μέσω της εφαρμογής αυτής προσπαθούν να συμβάλουν στην ακαδημαϊκή βιβλιογραφία προωθώντας τις προόδους στα συστήματα γεωγραφικών πληροφοριών (GIS). Προσπαθούν να αναπτύξουν μια αναλυτική προσέγγιση για την χώρο-χρονική προσβασιμότητα για κοινές δραστηριότητες και τι περιορισμοί μπαίνουν στην υλοποίηση αυτών, ενώ τα υπάρχοντα μοντέλα της χώρο-χρονικής προσβασιμότητας έχουν την τάση να επικεντρώνονται σε ατομικές δραστηριότητες ή/και να αναλύουν κοινές δραστηριότητες, εάν αυτές είναι ισοδύναμες με τις ατομικές δραστηριότητες. Θέλουν να συμβάλουν όχι μόνο στην ανάπτυξη ενός συστήματος γεωγραφικών πληροφοριών προσανατολισμένο στον άνθρωπο αλλά και στη διάδοση της ανάλυσης χώρο-χρονικής προσβασιμότητας σε ένα ευρύτερο κοινό. Τους ενδιαφέρει γενικά να απευθυνθούν σε διάφορων ειδικοτήτων επιστήμονες και να τους διαδώσουν το σκεπτικό των χωροχρονικών εφαρμογών και δυνατοτήτων επεξεργασίας τους. Γι αυτό η εφαρμογή είναι απλή, άμεση και πραγματικού χρόνου [NVS2010].

Στην εργασία αυτή δίνουν τις βασικές έννοιες του χωροχρονικού μονοπατιού (space - time path), χωροχρονικού πρίσματος (space - time prism) και της περιοχής δυνητικών μονοπατιών (potential path area) έχοντας υπόψη τη σημασία των περιορισμών για την ολοκλήρωση των ανθρωπίνων δραστηριοτήτων στο χώρο και το χρόνο. Ένα εννοιολογικό πλαίσιο για την ανάλυση της ανθρώπινης κίνησης και των δραστηριοτήτων σε ένα ολοκληρωμένο χωροχρονικό

περιβάλλον παρέχεται στην εργασία του Hagerstrand (1970). Ένα χωροχρονικό μονοπάτι (space - time path) είναι η αδιάκοπη σειρά των κινήσεων ενός ατόμου και σταθερών δραστηριοτήτων στο χώρο και το χρόνο. Επίσης, αναφέρεται στο πόσο επηρεάζει η ταχύτητα του αντικειμένου, το μονοπάτι. Όσο πιο γρήγορες είναι οι ταχύτητες κίνησης, τόσο πιο επίπεδα (με πολύ μικρές κλίσεις) είναι τα μονοπάτια στο διάγραμμα του χωροχρόνου. Δίνει τον ορισμό του χώρου δυνητικών μονοπατιών (Potential Path Space PPS) και οριοθετείται από ένα χωροχρονικό πρίσμα (space-time prism). Το χωροχρονικό πρίσμα είναι το σημείο τομής των δύο κώνων. Όλες οι παραπάνω έννοιες απεικονίζονται στην Εικόνα 2-9. Με αυτό το σχήμα, περιγράφεται η ελευθερία του ατόμου να κινείται αλλά και να λαμβάνει μέρος σε ευέλικτες δραστηριότητες, και έτσι τελικά να δίνει πληροφορίες για την διαθεσιμότητα του συγκεκριμένου ατόμου. Τέλος, η περιοχή πιθανών μονοπατιών (potential path area PPA) μπορεί να προκύψει από την προβολή ενός χωροχρονικού πρίσματος πάνω στο 2D γεωγραφικό επίπεδο. Απεικονίζει έτσι την χωρική έκταση που ένα άτομο μπορεί να καλύψει κατά τη διάρκεια του συνολικού χρόνου. Η καθημερινή περιοχή πιθανών μονοπατιών (daily potential path area DPPA) περιλαμβάνει το σύνολο των PPAς μεταξύ όλων των ζευγών των διαδοχικών δραστηριοτήτων που ένα άτομο πρέπει να διεξάγει κατά τη διάρκεια της ημέρας. Οι παραπάνω έννοιες μπορούν ευθέως να επεκταθούν σε πολλαπλά πρόσωπα που επιθυμούν να ξεκινήσουν μια κοινή δραστηριότητα. Ο χώρος δυνητικών μονοπατιών (PPS) μπορεί να επεκταθεί ως πιθανή χωροχρονική αλληλεπίδραση (potential interaction space-time PIS), η οποία ορίζεται ως το σύνολο των χωροχρονικών σημείων που όλα τα μέλη της ομάδας μπορούν να φτάσουν [NWW2007]. Η PIS κατασκευάζεται από τη τομή των PPSs των πολλαπλών ατόμων. Ομοίως, η δυνητική περιοχή αλληλεπίδρασης (potential interaction area PIA) μπορεί να οριστεί ως η προβολή της PIS στο γεωγραφικό επίπεδο (άξονες x και y). Η καθημερινή πιθανή περιοχή αλληλεπίδρασης (DPIA) δίνεται από το σύνολο των PIAs που είναι προσita σε πολλά άτομα για τη διεξαγωγή μιας κοινής δραστηριότητας επί διάστημα μιας ημέρας [NVS2010], [NWD2007].



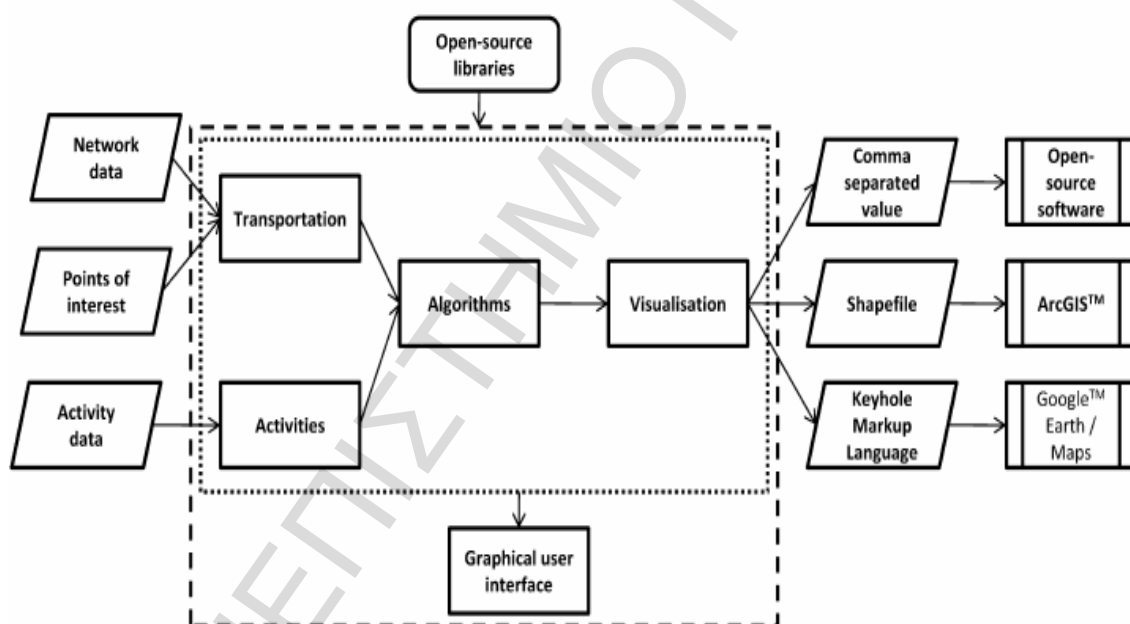
Εικόνα 2-9: Εννοιολογικό πλαίσιο

Στη συνέχεια ορίζονται κάποιες μετρικές, όπως το Number of Opportunities, (π.χ. ο αριθμός των fast-food που είναι προσπελάσιμος για μια ημέρα), όπου οι ευκαιρίες έχουν την ίδια βαρύτητα, το Activity Duration (π.χ. ο συνολικά διαθέσιμος χρόνος για ψώνια σε μια ημέρα) που όσο μεγαλύτερος είναι τόσο μεγαλύτερη ελευθερία δίνει στο άτομο για το πότε θα κάνει αυτή τη δραστηριότητα, το Locational Benefits είναι τα οφέλη που έχει από τις δυνατότητες που δίνει ένα μέρος όταν συμμετέχει σε μία δραστηριότητα που συντελείται σε αυτό. Στη συνέχεια γίνεται μια προσπάθεια να επεκταθούν αυτές οι μετρικές της ατομικής προσβασιμότητας σε μετρικές που

καλούνται από κοινού προσβασιμότητα. Δηλαδή να αξιολογηθούν οι δυνατότητες που είναι προσβάσιμες σε ομάδα ατόμων για από κοινού δραστηριότητα [NVS2010].

Geocomputational Toolkit: Η εφαρμογή αυτή είναι αυτόνομη και ανεξάρτητη από την πλατφόρμα που χρησιμοποιείται. Η Εικόνα 2-10 απεικονίζει την δομή της εφαρμογής. Πέντε αλληλένδετες ενότητες που έχουν υλοποιηθεί είναι: μεταφορές (transportation), δραστηριότητες (activities), αλγόριθμοι (algorithms), οπτικοποίηση (visualization) και γραφική διεπαφή χρήστη (graphical user interface GUI) [NVS2010].

Η μονάδα μεταφοράς βασίζεται σε χαρακτηριστικά που συνδέονται με τις ακμές και τους κόμβους του γραφήματος μεταφοράς και πληροφορίες σχετικά με σημεία ενδιαφέροντος (POI). Αυτά τα POIs είναι θέσεις όπου ένα ή περισσότερα άτομα μπορεί να προβούν σε συγκεκριμένη δραστηριότητα. Παραδείγματα POIs περιλαμβάνουν αθλητικές εγκαταστάσεις, καταστήματα, πρατήρια καυσίμων, βιβλιοθήκες, εστιατόρια, κλπ. Ένα σημείο ενδιαφέροντος (Point Of Interest) ή POI, είναι μια συγκεκριμένη θέση ενός σημείου που κάποιος μπορεί να το βρει χρήσιμο ή ενδιαφέρον. Για παράδειγμα στις περισσότερες εφαρμογές χρησιμοποιούν τον ορισμό αυτό όταν αναφέρονται σε σημεία όπως, εστιατόρια, πρατήρια καυσίμων, ξενοδοχεία, μουσεία, θέατρα, κάμπινγκ και πολλά άλλα σημεία ενδιαφέροντος ανάλογα με τις κατηγορίες που χρησιμοποιούνται στις εφαρμογές αυτές. Μια περιοχή ενδιαφέροντος (Region Of Interest) ή (ROI) έχει παρόμοια έννοια ή ιδέα, δηλώνοντας μια περιοχή η οποία μπορεί να περιέχει διάφορα σημεία ενδιαφέροντος POI. Ένα GPS σημείο ενδιαφέροντος προσδιορίζει, τουλάχιστον, το γεωγραφικό πλάτος και μήκος του POI, έχοντας υπόψη ένα συγκεκριμένο δεδομένο χάρτη. Οι GPS εφαρμογές συνήθως χρησιμοποιούν εικονίδια για να αντιπροσωπεύουν διαφορετικές κατηγορίες POI στον χάρτη.

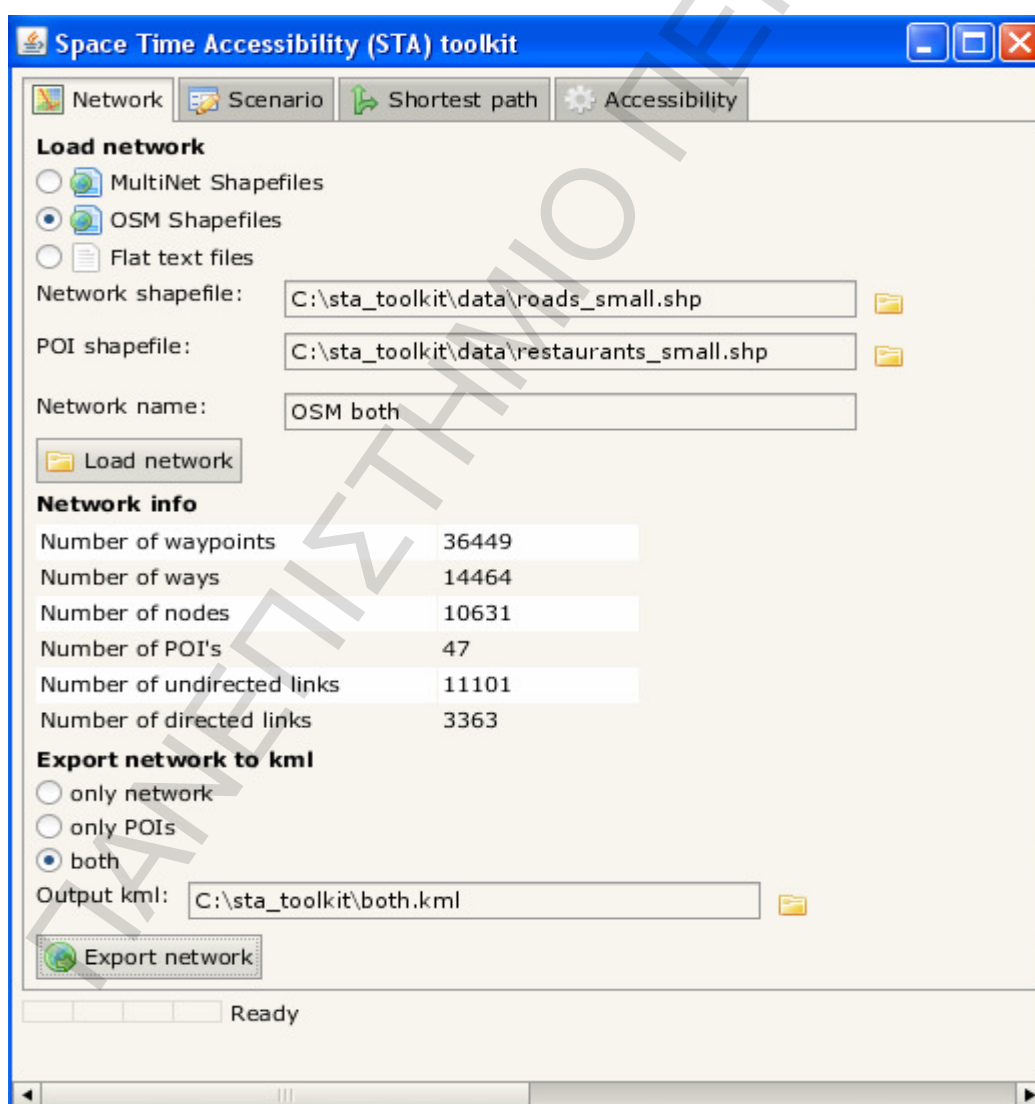


Εικόνα 2-10: Γενική δομή του σχεδιασμού

Η μονάδα των δραστηριοτήτων περιέχει ένα αντικείμενο με προσανατολισμό το μοντέλο των σταθερών δραστηριοτήτων των ατόμων που εμπλέκονται σε ένα συγκεκριμένο σενάριο. Η μονάδα αλγορίθμων αποτελείται από υποκατηγορίες, διότι ένα σενάριο μπορεί να αποτελείται από πολλαπλά πρόσωπα, το καθένα με το δικό του πρόγραμμα με σταθερές δραστηριότητες. Η ενότητα GUI επιτρέπει στο χρήστη να δημιουργήσει ένα σενάριο πολλαπλών ατόμων, να αποθηκεύσει ή να φορτώσει ένα συγκεκριμένο σενάριο. Και οι δύο ενότητες (δραστηριότητες και μεταφορά) χρησιμοποιούνται για να τροφοδοτήσουν την μονάδα αλγορίθμων. Τα αποτελέσματα μπορεί να εξαχθούν χρησιμοποιώντας τη μονάδα οπτικοποίησης σε Keyhole Markup Language (KML) αρχεία για την απεικόνιση στο Google Earth/Maps και σε Comma Separated Values (CSV) αρχεία για την απεικόνιση στην ESRI's ArcGIS και σε open-source GIS λογισμικό [NVS2010].

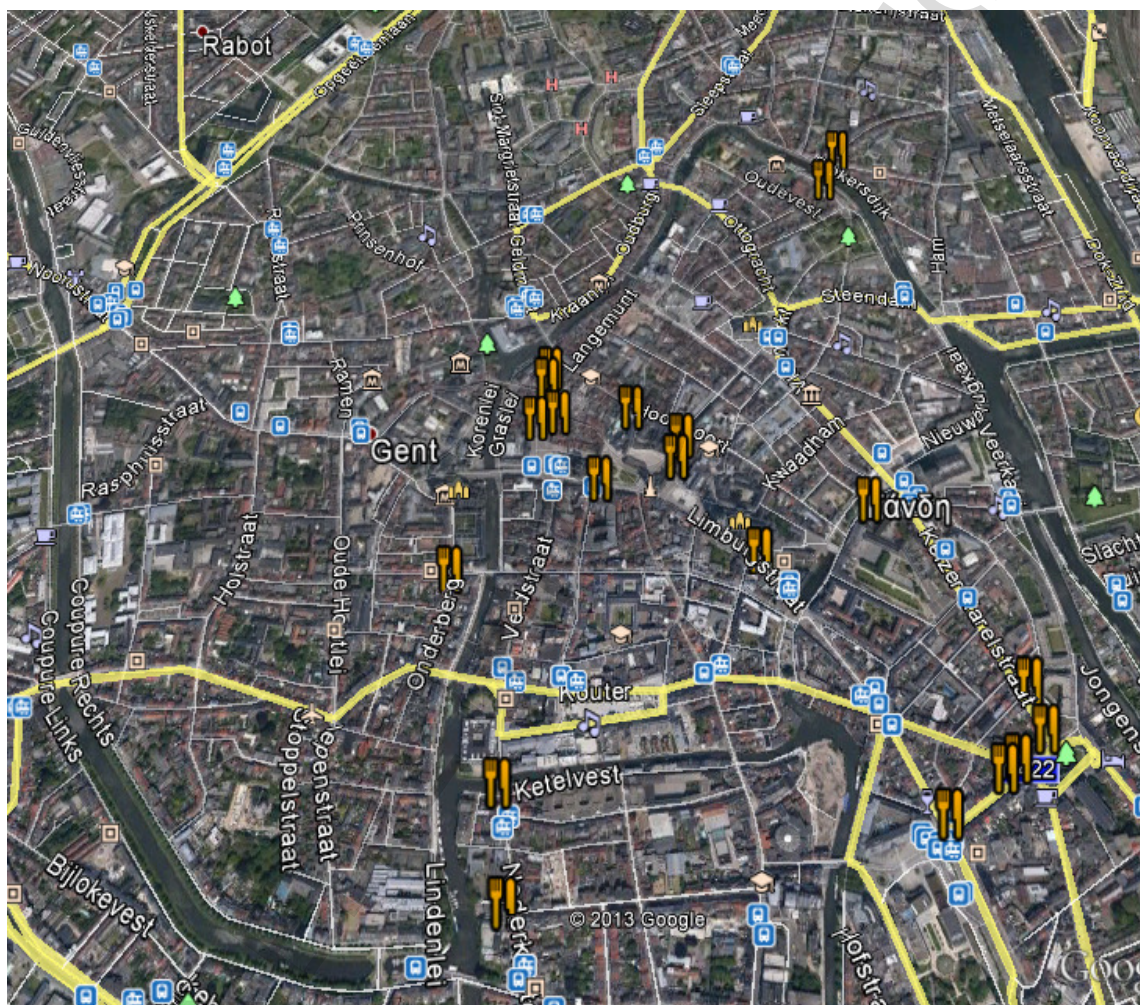
Η γραφική διεπαφή χρήστη χωρίζεται σε τέσσερα τμήματα: Δίκτυο, Σενάριο, Συντομότερου μονοπατιού, Προσβασιμότητα. Το τμήμα δικτύου είναι αυτό όπου τα δεδομένα θα πρέπει να φορτώνονται στο πρόγραμμα, και όπου μπορούμε να δούμε κάποια στατιστικά σχετικά με αυτά τα δεδομένα. Τρεις τύποι δεδομένων είναι απαραίτητοι: Πρώτον τα δεδομένα δικτύου μεταφοράς που περιγράφουν τις διαδρομές και τα σημεία τα οποία αντιπροσωπεύονται από ακμές και κόμβους ενός γραφήματος. Δεύτερον, τα δεδομένα με τα σημεία ενδιαφέροντος POIs, ο χρήστης μπορεί να ανοίξει ένα αρχείο με πληροφορίες για συγκεκριμένα σημεία ενδιαφέροντος. Τρίτον, πληροφορίες για την τοποθεσία και τις ώρες έναρξης και λήξης των σταθερών δραστηριοτήτων [URL5].

Το τμήμα σενάριο χρησιμοποιείται για να φορτώσουμε ένα σενάριο. Βασικά, είναι ένα αρχείο XML που περιέχει τα δεδομένα που σχετίζονται με τις σταθερές δραστηριότητες (τόπος, ώρα έναρξης, ώρα λήξης) μιας ομάδας ατόμων. Μπορούμε να δημιουργήσουμε ένα σενάριο από το μηδέν καθώς, και να το αποθηκεύσουμε για περαιτέρω χρήση. Στο τμήμα συντομότερου μονοπατιού, μπορούμε να υπολογίσουμε και να απεικονίσουμε μια μικρότερη / ταχύτερη διαδρομή. Στο τμήμα προσβασιμότητα μπορούμε να υπολογίσουμε μερικά πολύ ενδιαφέροντα χώρο-χρονικά μέτρα προσβασιμότητας. Πέντε διαφορετικά μέτρα προσβασιμότητας που μπορεί να υπολογίσει το πακέτο toolkit: DistanceAccessibility, TimeAccessibility, TimePrism, MultiTimePrism, Intersection [URL5].



Εικόνα 2-11: Η γραφική διεπαφή χρήστη

Παρακάτω στην Εικόνα 2-12 βλέπουμε ένα παράδειγμα. Η περιοχή μελέτης είναι η πόλη της Γάνδης, στο Βέλγιο. Το κέντρο της Γάνδης είναι η πιο πυκνοκατοικημένη περιοχή της πόλης και παρουσιάζει μια υψηλή συγκέντρωση από εστιατόρια. Στην παρακάτω εικόνα φαίνεται το οδικό δίκτυο και τα εστιατόρια που υπάρχουν στο κέντρο της πόλης. Στην εικόνα 2-11 στο τμήμα του δικτύου φορτώνουμε δυο shapefiles αρχεία, το ένα είναι το οδικό δίκτυο roads_small.shp και το άλλο αρχείο είναι τα σημεία ενδιαφέροντος restaurants_small.shp, δηλαδή, τα εστιατόρια στην περίπτωση μας. Πατώντας το κουμπί «Load network» το toolkit επιστρέφει κάποιες πληροφορίες σχετικά με το δίκτυο όπως το number of way, number of nodes, number of pois κλπ. Εξάγουμε το αποτέλεσμα είτε του δικτύου είτε των pois είτε και τα δυο μαζί σε ένα kml αρχείο, both.kml όπως φαίνεται στην Εικόνα 2-11 και πατώντας το κουμπί «Export network» μπορούμε να δούμε το αποτέλεσμα στο Google Earth, Εικόνα 2-12.



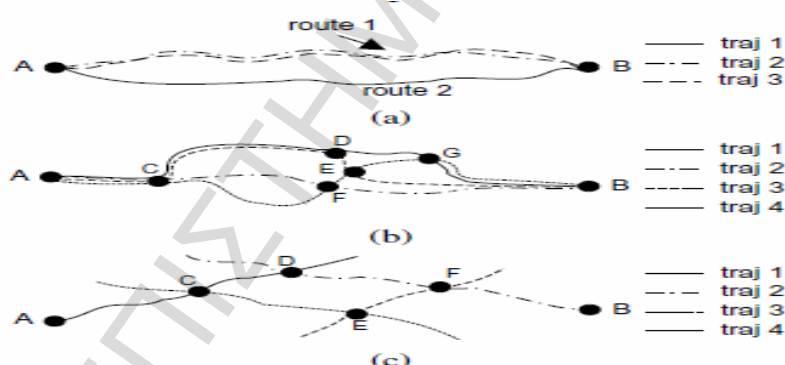
Εικόνα 2-12: Απεικόνιση εστιατορίων της πόλης της Γάνδης στο Google_Earth

2.8 Ανακάλυψη Δημοφιλής Διαδρομής από Ιστορικές Τροχιές

Η ενότητα αυτή, εστιάζει στην ανακάλυψη της πιο δημοφιλής διαδρομής (Most Popular Route) μεταξύ δύο θέσεων παρατηρώντας τις ταξιδιωτικές συμπεριφορές προηγούμενων χρηστών. Αυτό το ερώτημα είναι επωφελής για ταξιδιώτες οι οποίοι σχεδιάζουν ένα ταξίδι σε μια άγνωστη περιοχή, καθώς ιστορικές ταξιδιωτικές εμπειρίες μπορεί να αποκαλύψουν πώς οι άνθρωποι επιλέγουν συνήθως οδούς ανάμεσα στις τοποθεσίες. Η παρουσία κινητών συσκευών έχει δημιουργήσει ένα νέο φάσμα των location-based υπηρεσιών, οι οποίες γίνονται όλο και πιο δημοφιλής στις μέρες μας. Στο Google Maps, μπορούμε να απολαμβάνουμε την ευκολία των location - based υπηρεσιών, όπως ο σχεδιασμός οδηγών διαδρομών, εύρεση εστιατορίων, κλπ.

Μελετάμε το πρόβλημα του σχεδιασμού μιας ταξιδιωτικής διαδρομής εξετάζοντας τις ιστορικές τροχιές άλλων ανθρώπων που παράγονται από GPS συσκευές. Μια συλλογή από τροχιές δίνει συμβουλές για το πώς οι άνθρωποι συνήθως ταξιδεύουν μεταξύ τοποθεσιών, και στόχος είναι η ανακάλυψη της πιο δημοφιλής διαδρομής από μια συγκεκριμένη τοποθεσία σε άλλη. Αυτό είναι εντελώς διαφορετικό από τις υπάρχουσες μεθόδους συντομότερης διαδρομής. Η πιο δημοφιλής διαδρομή είναι στατιστικό αποτέλεσμα που προέρχεται από πραγματικές ταξιδιωτικές διαδρομές που διεξάγονται από άλλους ανθρώπους και δεν είναι κατ' ανάγκη η συντομότερη διαδρομή. Για παράδειγμα, τουρίστες που ταξιδεύουν σε εθνικό πάρκο πιθανόν ακολουθήσουν μια διαδρομή από την είσοδο μέχρι την έξοδο που καλύπτει τα περισσότερα από τα σημεία ενδιαφέροντος, αν ακολουθούσαν το συντομότερο μονοπάτι μπορεί να χάσουν πολλά αξιοθέατα.

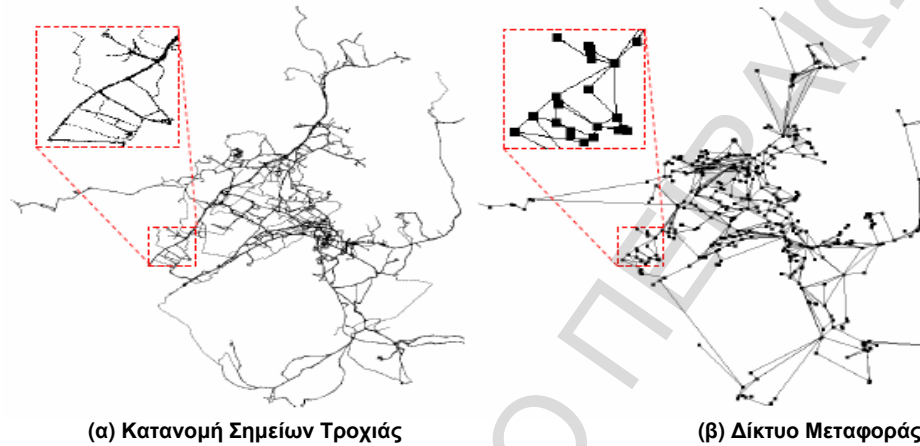
Με δεδομένη την αφετηρία και τον προορισμό μιας τοποθεσίας, μπορεί κανείς να ελέγξει όλες τις υπάρχουσες διαδρομές που συνδέουν τα δύο σημεία και να μετρήσει τον αριθμό των τροχιών μέσα από κάθε μία από τις διαδρομές. Η διαδρομή με τη μεγαλύτερη στήριξη υποτίθεται ότι είναι η πιο δημοφιλής. Για παράδειγμα, στην Εικόνα 2-13(α), υπάρχουν 2 τροχιές (traj 2 & 3) που περνούν την διαδρομή 1 από τη θέση A στη B, ενώ μόνο μία τροχιά (traj 1) στη διαδρομή 2, οπότε θα λέγαμε η διαδρομή 1 είναι προτιμότερη. Συνήθως δεν είναι δυνατή η εύρεση από τέτοιες καλές ομάδες από τροχιές. Στην Εικόνα 2-13(β), 4 τροχιές (traj 1-4) ενώνουν την τοποθεσία A και B. Όλες οι τροχιές τέμνονται και θα μπορούσαν να υπάρχουν πολλές πιθανές διαδρομές (π.χ., ACFB, AC FEB, κ.λπ.). Στην προκειμένη περίπτωση, μια συγκεκριμένη και λογική συνάρτηση δημοτικότητας είναι απαραίτητο να μετρήσει πόσο δημοφιλής είναι η διαδρομή. Ο όρος «δημοφιλής» είναι υποκειμενικός. Σε περίπτωση που δεν υπάρχει τροχιά που συνδέει τα A και B άμεσα, Εικόνα 2-13(γ), η πρόταση μιας συνδυασμένης διαδρομής θα είναι πιο χρήσιμη για τους χρήστες. Υποθέτουμε ότι ACD, Εικόνα 2-13(ε), είναι μία ακολουθία από περιοχές με υψηλή στήριξη. Ωστόσο, οι άνθρωποι που ακολουθούν την ACD θα μπορούσαν να πάνε σε οποιαδήποτε θέση αντί του προορισμού B. Πιθανόν οι άνθρωποι να πηγαίνουν στο B μέσω ACEF, έστω και αν η στήριξη της ACEF δεν είναι υψηλή. Η ACD δεν είναι ακριβής για να αντικατοπτρίζει συμπεριφορές χρηστών σε σχέση με το B προορισμό [CSZ2011].



Εικόνα 2-13: Παραδείγματα ανακάλυψης δημοφιλών διαδρομών

Η μέτρηση του αριθμού τροχιών δεν είναι αρκετό για την ανακάλυψη της δημοφιλής διαδρομής ανάμεσα σε δύο τοποθεσίες, λόγω του μεγάλου αριθμού των πιθανών διαδρομών και της δυσκολίας συνδυασμού τμημάτων τροχιάς. Βασική ιδέα είναι η κατασκευή ενός δικτύου μεταφοράς από τις πρώτες τροχιές. Κάθε κόμβος σε ένα δίκτυο μεταφοράς θεωρείται ως μία «σημαντική θέση». Αντλούμε την πιθανότητα της μεταφοράς από κάθε «σημαντική θέση» στον τόπο προορισμού με βάση τις ιστορικές τροχιές, και η πιθανότητα μεταφοράς χρησιμοποιείται ως δείκτης της δημοτικότητας. Στη συνέχεια, η δημοτικότητα μιας διαδρομής προς τον προορισμό ορίζεται ως το γινόμενο των πιθανοτήτων μεταβίβασης όλων των «σημαντικών θέσεων» στη διαδρομή. Για την επίτευξη αυτού του στόχου, έχει αναπτυχθεί αρχικά ο *Coherence Expanding* αλγόριθμος για να ανακτήσει ένα δίκτυο μεταφοράς. Μετά, εφαρμόζεται το *Absorbing Markov Chain* μοντέλο για να εντοπίσει μια λογική πιθανότητα μεταφοράς για κάθε κόμβο, η οποία χρησιμοποιείται ως δείκτης δημοτικότητας. Τέλος, ο *Maximum Probability Product* αλγόριθμος βρίσκει την πιο δημοφιλή διαδρομή από ένα δίκτυο μεταφοράς με βάση τους δείκτες δημοτικότητας. Ο αλγόριθμος μοιράζεται το ίδιο πνεύμα με τον αλγόριθμο Dijkstra [CSZ2011].

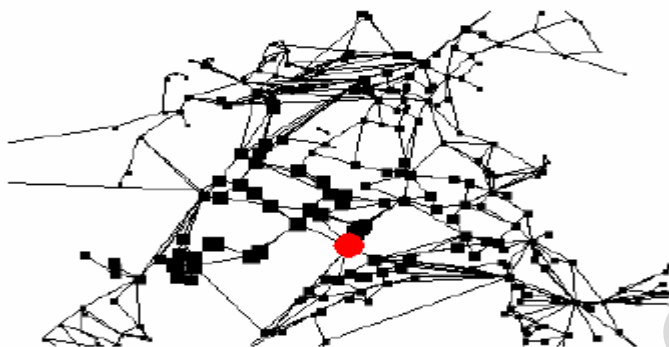
Εξόρυξη Δικτύου Μεταφοράς: Το δίκτυο μεταφοράς είναι ένα διάγραμμα $G(N,E)$ που αναφέρει τις κινήσεις ανάμεσα στις τοποθεσίες. N είναι ένα σύνολο κόμβων μεταφοράς, το οποίο μπορεί να είναι διασταύρωση τροχιών ή απλά το τέλος τροχιάς. E είναι μια συλλογή από ακμές που συνδέουν τους κόμβους μεταφοράς. Υφίσταται μια ακμή e από τον κόμβο A στο B εάν υπάρχει τουλάχιστον μία συνεχόμενη τροχιά από A στο B χωρίς άλλους κόμβους μεταφοράς ενδιάμεσα. Για τις τροχιές που κινούνται προς την ίδια κατεύθυνση μεταξύ δύο γειτονικών κόμβων, μπορεί να ομαδοποιούνται στο ίδιο άκρο. Κατά συνέπεια, οι τροχιές μετατρέπονται σε ένα δίκτυο δρομολόγησης. Όπως απεικονίζεται στην Εικόνα 2-14, στόχος είναι η απόκτηση του δικτύου μεταφοράς, Εικόνα 2-14(β), από ένα σύνολο τροχιών, Εικόνα 2-14(α). Στην Εικόνα 2-14(b), μια κουκίδα αντιπροσωπεύει έναν κόμβο και η γραμμή μια ακμή μεταφοράς.



Εικόνα 2-14: Εξόρυξη Δικτύου Μεταφοράς

Στον αλγόριθμο Coherence Expanding, ελέγχεται απλά κάθε σημείο της τροχιάς στο P διαδοχικά. Αν δεν έχει καταταχτεί σε κανένα cluster ακόμα, προσπαθεί να το επεκτείνει με τη χρήση του Directly Coherence-Connected. Μετά από αυτό, αν το μέγεθος του επιστρεφόμενου συνόλου των σημείων υπερβαίνει ή είναι ίσο με το κατώφλι ϕ , τότε το σύνολο αποθηκεύεται ως έγκυρο cluster. Με αυτόν τον τρόπο, τελικά, όλες οι έγκυρες ομάδες θα βρεθούν, από τη στιγμή που θα ξεκινήσει τον έλεγχο σε οποιοδήποτε σημείο της συστάδας, όλα τα άλλα σημεία του cluster θα ανακτηθούν στην διαδικασία της επέκτασης. Για τα σημεία που δεν ανήκουν σε καμία έγκυρη συστάδα, απλώς θα τα παραλείψει [CSZ2011].

Εξαγωγή Πιθανότητας Μεταφοράς: Στην ενότητα αυτή πραγματοποιείται η ανάλυση ταξιδιωτικών συμπεριφορών χρηστών σε ένα δίκτυο και η εξαγωγή πιθανοτήτων μεταφοράς σε δεδομένο προορισμό. Στόχος είναι να μάθουμε ποιος κόμβος είναι πιο πιθανό να οδηγήσει το χρήστη στο προορισμό και η πιθανότητα θα χρησιμεύσει ως δείκτης δημοτικότητας. Σε έναν κόμβο μεταφοράς n_i , ένας τρόπος για την παρατήρηση ιστορικών συμπεριφορών χρηστών είναι να απαριθμήσει όλες τις γειτονικές ακμές που ξεκινούν από n_i και να ελέγξει πόσοι άνθρωποι πέρασαν από αυτές. Αυτά τα στατιστικά στοιχεία της συμπεριφοράς χρηστών είναι μια γενική περίπτωση χωρίς την μελέτη του προορισμού. Ως εκ τούτου, σε ένα κόμβο ενός συγκεκριμένου προορισμού θα πρέπει να εξεταστεί εάν οι ιστορικές τροχιές είναι με κατεύθυνση τον προορισμό ή όχι. Σε δίκτυο που βασίζεται στην πιθανότητα στροφής σε ένα κατευθυνόμενο γράφημα με την πιθανότητα μετάβασης από τον κόμβο n_i στο n_j να είναι $Pr_d(n_i \rightarrow n_j)$. Καλείται πιθανότητα μεταφοράς και λαμβάνει υπόψη t ακόλουθες μεταφορές. Έτσι, εξετάζονται όλες οι πιθανές ακμές σε t βήματα. Όσο μεγαλύτερη πιθανότητα μεταφοράς ένας κόμβος n_i κατέχει, τόσο μεγαλύτερη πιθανότητα ότι n_i θα οδηγήσει στον προορισμό. Η πιθανότητα μεταφοράς $Pr^t(n_i \rightarrow d)$ της μετάβασης από n_i στον προορισμό d σε t βημάτων είναι το άθροισμα των πιθανοτήτων στο d σε $1, 2, \dots, t$ βήματα. Η πιθανότητα μεταφοράς μπορεί να χρησιμοποιηθεί ως δείκτης για το πόσο δημοφιλής είναι ο κόμβος n_i στον προορισμό d . Στην Εικόνα 2-15, ζωγραφίζονται οι κόμβοι μεταφοράς από ορθογώνια με το μέγεθος ανάλογη με την πιθανότητα μεταφοράς. Ο προορισμός εμφανίζεται ως ένας κύκλος. Αν $t = 20$, βλέπουμε ότι οι περισσότεροι άνθρωποι ταξιδεύουν στον προορισμό μέσω των κόμβων στην αριστερή πλευρά (μεγαλύτερα ορθογώνια).



Εικόνα 2-15: Διανομή Πιθανότητας Μεταφοράς

Για τον υπολογισμό της πιθανότητας μεταφοράς για όλους τους κόμβους σε ένα δίκτυο μεταφοράς, υιοθετήθηκε το μοντέλο Absorbing Markov Chain, το οποίο είναι ένας ειδικός τύπος του Markov Chain με τουλάχιστον μία *απορροφητική* (absorbing) κατάσταση. Μια κατάσταση (κόμβος) n_i μιας αλυσίδας Markov ονομάζεται απορρόφηση εάν η πιθανότητα μετάβασης από n_i στο n_i (η ίδια) είναι πάντα 1, ενώ οι μη απορροφητικές καταστάσεις ονομάζονται παροδικές καταστάσεις. Σε κατευθυνόμενο δίκτυο ο κόμβος προορισμού αντιμετωπίζεται ως απορροφητική κατάσταση. Τα τελικά σημεία της τροχιάς χωρίς εξερχόμενες ακμές θεωρούνται επίσης ως απορροφητικές καταστάσεις διότι δεν μπορεί να κινηθεί σε άλλο κόμβο σε ένα κατευθυνόμενο δίκτυο. Όλοι οι άλλοι κόμβοι μεταφοράς θεωρούνται παροδικές καταστάσεις. Η απορρόφηση των καταστάσεων που μεταφέρουν στους εαυτούς τους έχει πιθανότητα 1, ενώ οι παροδικές καταστάσεις κάνουν τη μετάβαση σε γειτονικούς κόμβους, σύμφωνα με τη πιθανότητα στροφής.

Ο αλγόριθμος Deriving Transfer Probability απαριθμεί τις διαδικασίες για την εξαγωγή των πιθανοτήτων μεταφοράς για ένα δίκτυο μεταφοράς $G(N, E)$. Η επιλογή μιας κατάλληλης t είναι επίσης σημαντικό για την παραγωγή των πιθανοτήτων μεταφοράς. Προσδιορίζει το μέγιστο βήμα και το μήκος της μακρύτερης διαδρομής. Αν το μήκος μιας διαδρομής είναι μεγάλο, δεν έχει νόημα καθώς οι άνθρωποι δεν θα ταξιδέψουν σε τέτοια διαδρομή. Εάν t είναι μικρό, για παράδειγμα, ακόμη μικρότερο από τον αριθμό βημάτων της συντομότερης διαδρομής προς τον προορισμό, τότε θα αποτύχουμε να ανακαλύψουμε μια διαδρομή, επειδή δεν υπάρχει διαδρομή που μπορεί να φτάσει στο προορισμό μέσα σε t βήματα (δηλαδή, η πιθανότητα μεταφοράς = 0). Λαμβάνοντας υπόψη τους δύο παράγοντες, θέτουμε t ως τη διάμετρο του δικτύου, που εγγυάται τουλάχιστον ότι μία διαδρομή μπορεί να βρεθεί μεταξύ δύο κόμβων [CSZ2011].

Αναζήτηση της πιο Δημοφιλής Διαδρομής: Με την εξόρυξη του δικτύου μεταφοράς και την παραγωγή πιθανοτήτων μεταφοράς, αποκτάμε ένα κατευθυνόμενο δίκτυο μεταφοράς με ένα σύνολο πιθανοτήτων μεταφοράς υποδεικνύοντας πώς ένας κόμβος μεταφοράς να οδηγήσει στο d προορισμό. Παίρνουμε την πιθανότητα μεταφοράς ενός κόμβου n_i και d ο δείκτη δημοτικότητας. Αν $n_i = d$, υποθέτουμε $n_i.popularity(d)=1$, και αν n_i είναι ένα τελικό σημείο τροχιάς που ανήκει στις παροδικές καταστάσεις, θέτουμε $n_i.popularity(d)=0$. Κάθε κόμβος n_i διατηρεί m δείκτες: $n_i.popularity(n_1), \dots, n_i.popularity(n_m)$, για όλους τους κόμβους m στο δίκτυο που είναι πιθανοί προορισμοί. Ένας δείκτης εκφράζει τη δημοτικότητα του κόμβου για να πάει κανείς στο αντίστοιχο προορισμό. Μελετάμε την ανακάλυψη της δημοφιλής διαδρομής σύμφωνα με τους δείκτες δημοτικότητας. Μια R διαδρομή ορίζεται ως μια ακολουθία κόμβων $n_1 \rightarrow n_2 \rightarrow \dots \rightarrow n_i$, όπου (n_i, n_{j+1}) , $(1 \leq j < i)$, είναι μια υπάρχουσα ακμή. Η δημοτικότητα $\rho(R)$ μιας διαδρομής όσον αφορά ένα d προορισμό ορίζεται ως το γινόμενο του δείκτη δημοτικότητας του κάθε κόμβου μεταφοράς.

Η πιο δημοφιλής διαδρομή από έναν κόμβο αρχή s σε ένα κόμβο προορισμού d είναι η διαδρομή $R = n_1 \rightarrow n_2 \rightarrow \dots \rightarrow n_i$, $(n_1=s, n_i=d)$, έτσι ώστε η τιμή $\rho(R)$ να μεγιστοποιείται ανάμεσα σε όλες τις πιθανές διαδρομές από s σε d . Στο MPP αλγόριθμο για την ανακάλυψη της MPR, καταγράφουμε τη μέγιστη $\rho()$ τιμή της διαδρομής από τον κόμβο s στον n_i σε μια ετικέτα $L(n_i)$, η οποία είναι αρχικοποιημένη στο 0 και μόνο το $L(s)$ έχει οριστεί να είναι 1 (γραμμή 1-2). Μια max ουρά προτεραιότητας PQ χρησιμοποιείται για να καθορίσει τον κόμβο με τη μέγιστη τιμή $\rho()$ ετικέτας από μη-σαρωμένους κόμβους. Στην αρχή, όλοι οι κόμβοι είναι μη-σαρωμένοι, έτσι SN είναι null και PQ περιέχει τον κόμβο s . Στο βρόχο while (γραμμή 6), εξάγουμε τον κόμβο u με τη μέγιστη ετικέτα από PQ και ενημερώνουμε τις ετικέτες γειτονικών κόμβων, (u,v) είναι μια υπάρχουσα ακμή μεταφοράς, (γραμμή 11-15). Αν $L(v) < L(u) \times v.popularity(d)$, που σημαίνει ότι

έχουμε βρει μια πιο δημοφιλής διαδρομή στο v μέσω u , τότε θα ενημερώσουμε την ετικέτα v και λαμβάνει u ως προκάτοχο του v στη διαδρομή. Εκτός αυτού, όλοι οι κόμβοι που ανακαλύψαμε προστίθενται στην ουρά προτεραιότητας για περαιτέρω εξέταση. Μόλις ο προορισμός d πεταχτεί έξω από την ουρά (γραμμή 8), η πιο δημοφιλής διαδρομή από s στο d ανακαλύφθηκε και ανακτούμε όλη τη διαδρομή ακολουθώντας το σύνδεσμο προκατόχου του κάθε κόμβου από d . Η πολυπλοκότητα είναι ίδια με εκείνη του Dijkstra, το οποίο είναι $O(|E| + |N|\log|N|)$, όπου $|E|$ ο αριθμός ακμών και $|N|$ ο αριθμός κόμβων. Η απόδειξη ορθότητας του αλγορίθμου είναι από μια μέθοδος επαγωγικής, που είναι παρόμοια με την απόδειξη του Dijkstra [CSZ2011].

Algorithm 4: Maximum Probability Product

input : A transfer network $G(N, E)$,
 $N = \{n_1, n_2, \dots, n_m\}$;
 Start node s ; Destination node d

output: The most popular route MPR

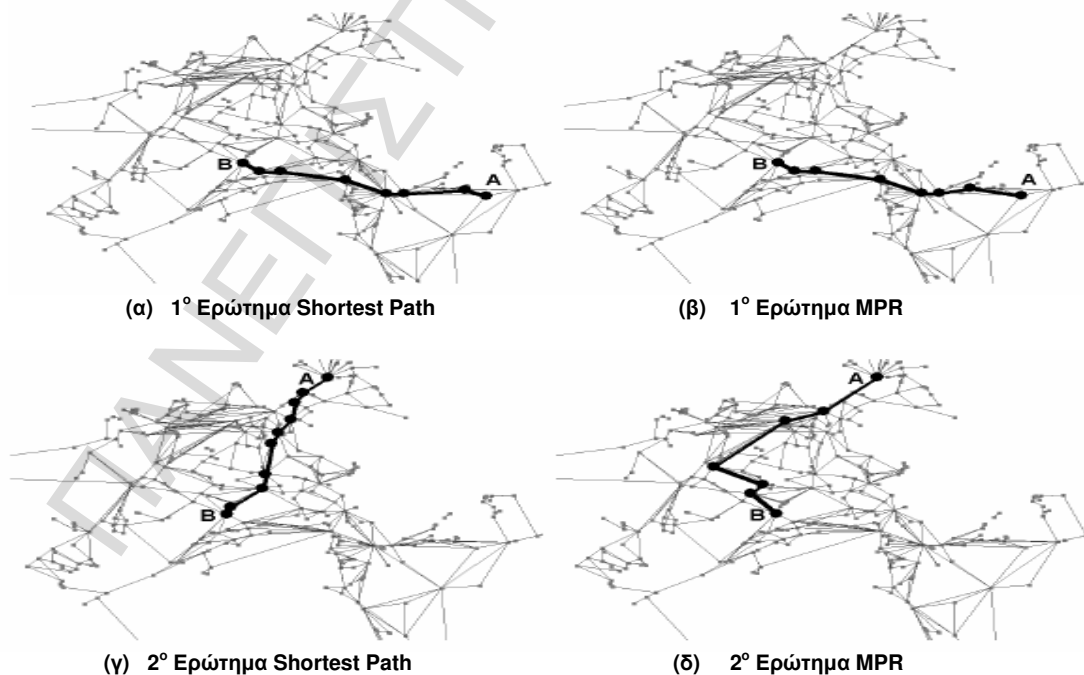
```

1 For all  $n_i \in N$ , label  $L(n_i) \leftarrow 0$ ;
2  $L(s) \leftarrow 1$ ;
3 Priority Queue  $PQ \leftarrow null$ ;
4 Scanned Nodes  $SN \leftarrow null$ ;
5  $PQ.enqueue(s)$ ;
6 while  $PQ \neq null$  do
7    $u \leftarrow PQ.extractMax()$ ;
8   if  $u = d$  then
9     return MPR;
10   $SN.add(u)$ ;
11  for each  $v \in u.adjacentNodes$  do
12    if  $L(v) < L(u) \times v.popularity(d)$  then
13       $L(v) \leftarrow L(u) \times v.popularity(d)$ ;
14       $v.predecessor \leftarrow u$ ;
15       $PQ.add(v)$ ;

```

Εικόνα 2-16: Maximum Probability Product Αλγόριθμος

Παραδείγματα MPR vs. Shortest Path: Τα αποτελέσματα αναζήτησης του MPR θα συγκρίνουν με τα αντίστοιχα αποτελέσματα του SP χρησιμοποιώντας παραδείγματα.



Εικόνα 2-17: Παραδείγματα Ερωτημάτων (A-Start Node, B-End Node)

Στο πρώτο παράδειγμα ερωτήματος, η πιο δημοφιλής διαδρομή (Εικόνα 2-17(β)) είναι σχεδόν η ίδια με την συντομότερη διαδρομή (Εικόνα 2-17(α)), όπου ο προορισμός σχεδιάζεται ως ένα ορθογώνιο. Ακόμα κι αν η MPR και η αντίστοιχη συντομότερη διαδρομή είναι σχεδόν η ίδια με την πρώτη ερώτηση, μπορούμε να βρούμε πολλά σενάρια που η MPR είναι πολύ διαφορετική από τη συντομότερη διαδρομή. Ένα παράδειγμα είναι στην Εικόνα 2-17(δ) όταν η MPR αποκλίνει προς τα αριστερά, ενώ η συντομότερη διαδρομή πηγαίνει κατ'ευθείαν κάτω στον προορισμό στην Εικόνα 2-17(γ). Το παράδειγμα αυτό αποδεικνύει ότι μια προτιμότερη MPR δεν είναι απαραίτητα να είναι η συντομότερη διαδρομή.

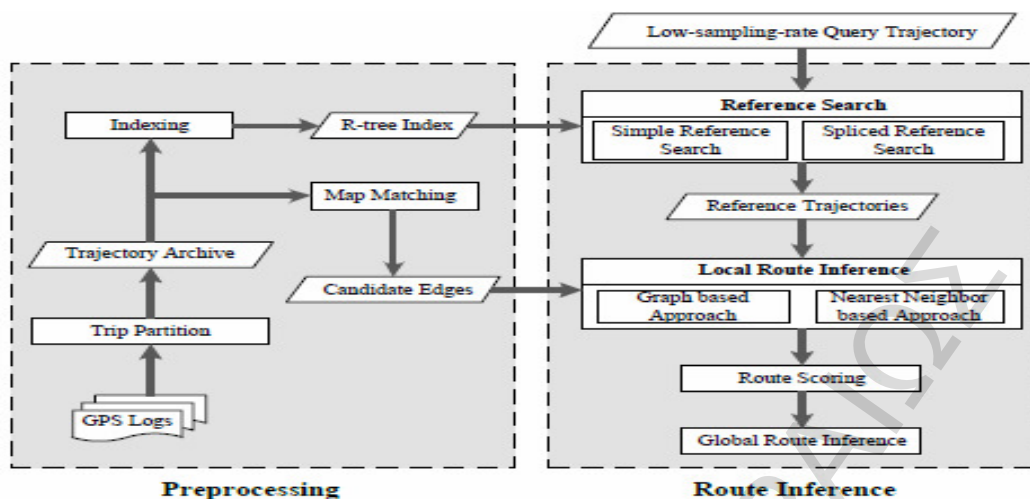
2.9 Μείωση Αβεβαιότητας Χαμηλού Ρυθμού Δειγματοληψίας σε Τροχιές

Η αυξανόμενη διαθεσιμότητα GPS ενσωματωμένο σε κινητές συσκευές έχει δημιουργήσει ένα νέο φάσμα των location-based υπηρεσιών, οι οποίες έχουν συσσωρεύσει μια τεράστια συλλογή τροχιών. Μια τροχιά τοποθεσίας είναι μια καταγραφή της διαδρομής από κινούμενα αντικείμενα, όπως είναι τα άτομα, τα οχήματα, τα ζώα και τα φαινόμενα της φύσης. Στην πράξη, ένα μεγάλο μέρος αυτών των τροχιών είναι χαμηλού ρυθμού δειγματοληψίας. Το χρονικό διάστημα μεταξύ διαδοχικών σημείων GPS κάποιων τροχιών μπορεί να είναι αρκετά λεπτά ή και ώρες. Με ένα τόσο χαμηλό ποσοστό δειγματοληψίας, τα περισσότερα στοιχεία της κίνησης χάθηκαν, το οποίο καθιστά δύσκολη την αποτελεσματική επεξεργασία. Για παράδειγμα, τα περισσότερα ταξί στις μεγάλες πόλεις είναι εξοπλισμένα με GPS, που τους επιτρέπουν να αναφέρουν χρονοσήμανση θέσεων σε ένα κέντρο δεδομένων περιοδικά. Λόγω του κόστους επικοινωνίας, τα ταξί αναφέρουν τις θέσεις σε χαμηλή συχνότητα. Η διάδοση GPS συσκευών έχει οδηγήσει σε ώθηση των location-based υπηρεσιών, όπου οι χρήστες μπορούν να αποκτήσουν τις παρούσες θέσεις τους, να αναζητήσουν ενδιαφέροντα μέρη και να βρουν την διαδρομή προς έναν προορισμό. Αυτές οι υπηρεσίες έχουν συσσωρεύσει τεράστιο ποσό δεδομένων τροχιάς θέσης και εμπνέει πολλές εφαρμογές, όπως η ανάλυση της ροής κυκλοφορίας, να αξιοποιήσουν αυτή την πληροφόρηση για την επίτευξη καλύτερης ποιότητας υπηρεσιών.

Με ένα τόσο χαμηλό ποσοστό δειγματοληψίας, λεπτομερείς πληροφορίες σχετικά με την ακριβή κίνηση των αντικειμένων χάνεται και μεγάλη αβεβαιότητα προκύπτει στις διαδρομές. Σε τροχιές, όπως αυτό από τα μεταναστευτικά πουλιά, όπου τα αρχεία τοποθεσιών είναι πολύ αραιά, η αβεβαιότητα γίνεται ακόμα πιο σημαντική. Αυτό το είδος της αβεβαιότητας θα επηρεάσει σοβαρά την αποτελεσματικότητα και την αποδοτικότητα της μεταγενέστερης διαδικασίας, όπως η ευρετηρίαση, αναζήτηση και εξόρυξη. Τα αποτελέσματα της μεθοδολογίας μπορεί να είναι επωφελής για πολλές εφαρμογές, όπως η διαχείριση της κυκλοφορίας, η εξόρυξη από geo-tag φωτογραφίες, η επιστημονική μελέτη της συμπεριφοράς των ζώων ή η κίνηση τυφώνων κλπ. Η μεθοδολογία που υιοθετήθηκε είναι να επωφεληθούν από τις ιστορικές τροχιές. Για μια πρακτική λύση, γίνεται η επεξεργασία ενός συγκεκριμένου ερωτήματος σε τρία βήματα. Πρώτον, διαιρείται το σύνολο ερωτημάτων σε μια σειρά από επιμέρους ερωτήματα και αναζητούμε για τις τροχιές αναφοράς. Μετά, θα συμπεράνουμε τα τοπικά δρομολόγια για κάθε δευτερεύον ερώτημα εξετάζοντας τις τροχιές αναφοράς. Τέλος, συνδέουμε τα τοπικά δρομολόγια για να σχηματίσουν τις καθολικές διαδρομές και να επιστρέψουν αυτά με την υψηλότερη βαθμολογία.

Για να μειωθεί η αβεβαιότητα δίνεται μια τροχιά χαμηλής δειγματοληψίας, στόχος είναι να συμπεράνουμε πιθανές διαδρομές. Η μεθοδολογία που ακολουθήθηκε είναι να επωφεληθούμε από τις ιστορικές τροχιές. Η λύση, History based Route Inference System (HRIS), καλύπτει μια σειρά νέων αλγορίθμων που μπορεί να αντλήσει το πρότυπο διαδρομής από ιστορικά στοιχεία και να την εντάξει στη διαδικασία εξαγωγής συμπερασμάτων διαδρομής. Τα αποτελέσματα δείχνουν ότι η HRIS μπορεί να επιτύχει μεγαλύτερη ακρίβεια από τους υπάρχοντες map-matching αλγορίθμους για τροχιές με χαμηλό ρυθμό δειγματοληψίας [ZZXZ2012].

Επισκόπηση Συστήματος: Στην Εικόνα 2-18 φαίνεται η αρχιτεκτονική του συστήματος, που αποτελείται από δύο μέρη: *προεπεξεργασία και συμπέρασμα διαδρομής*. Το πρώτο μέρος λειτουργεί offline και πρέπει να εκτελείται μία φορά, εκτός αν το αρχείο τροχιών ή το οδικό δίκτυο έχει ενημερωθεί. Το δεύτερο μέρος είναι το βασικό μέρος του συστήματος και θα πρέπει να διεξαχθεί online με βάση το ερώτημα που ορίζεται από ένα χρήστη [ZZXZ2012].



Εικόνα 2-18: Επισκόπηση Συστήματος

Προεπεξεργασία (Preprocessing): Για τη βελτίωση της ποιότητας των αρχείων τροχιών και για την διευκόλυνση της ακολουθούσας διαδικασίας, εκτελείται κάποια προεπεξεργασία πριν χρησιμοποιηθεί από το μέρος συμπίεσης διαδρομής.

Trip Partition. Ένα αρχείο καταγραφής GPS μπορεί να καταγράψει την κίνηση ενός αντικείμενου για μεγάλο χρονικό διάστημα, το οποίο θα ταξιδέψει από πολλές πηγές σε πολλούς προορισμούς. Εάν το αντικείμενο είναι ακίνητο για ένα σχετικά μεγάλο χρονικό διάστημα, συνεπάγεται το τέλος διαδρομής καθώς και την έναρξη της επόμενης διαδρομής. Για κάθε ιστορική τροχιά, θα εντοπίσει τα σημεία παραμονής και θα αφαιρέσει το σημείο GPS που ανήκει στο σημείο αυτό, μετά η τροχιά θα χωρίζεται σε διάφορες διαδρομές.

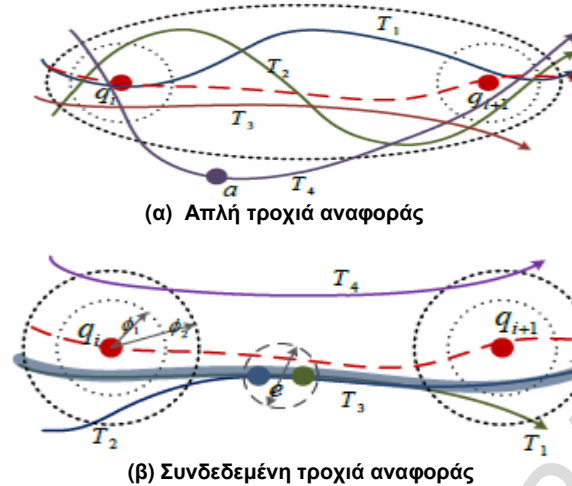
Map Matching. Λόγω λάθους μέτρησης, η παρατήρηση ενός GPS σημείου δεν αντικατοπτρίζει την πραγματική θέση του με ακρίβεια. Σε αυτό το βήμα, θα ευθυγραμμίζονται τα παρατηρούμενα GPS σημεία επί των τμημάτων του δρόμου με την εφαρμογή τεχνικών χάρτη αντιστοίχισης.

Indexing. Μια πραγματική βάση δεδομένων κινούμενων αντικειμένων αρχειοθετεί χιλιάδες τροχιές και εκατομμύρια GPS σημεία, που απαγορεύει να τα ψάξουμε διαδοχικά. Για να βελτιωθεί η απόδοση συστήματος, χρησιμοποιεί R-δέντρο για να οργανώσει όλα τα GPS σημεία.

Συμπέρασμα Διαδρομής (Route Inference): Αυτό είναι το βασικό μέρος του συστήματος, το οποίο θα επεξεργαστεί ένα συγκεκριμένο ερώτημα σε τρεις φάσεις :

Αναζήτηση Τροχιών Αναφοράς (Reference Trajectory Search). Χρησιμοποιείται η έννοια τροχιά αναφοράς για να περιγράψει τις ιστορικές τροχιές που είναι χρήσιμες για να βρεθούν οι πιθανές διαδρομές του ερωτήματος. Στην ιδανική περίπτωση, οι τροχιές αναφοράς θα πρέπει να ταξιδεύουν στην ίδια πορεία με το ερώτημα. Όχι όλες οι ιστορικές τροχιές είναι σχετικές με το ερώτημα λόγω αραιότητας δεδομένων. Προτείνεται να χωρίσει το ερώτημα σε ζεύγη θέσεων και αναζητώντας για τροχιές αναφοράς σε κάθε περιοχή [ZZXZ2012].

Simple Reference Trajectory (Απλή Τροχιά Αναφοράς). Η απλή τροχιά αναφοράς ορίζεται ως εξής: $A_s \text{ nh}(q, T)$ χαρακτηρίζει το πλησιέστερο σημείο της τροχιάς T σε σχέση με το σημείο q . Έχοντας υπόψη μια ακτίνα $\varphi > 0$, η τροχιά T_k είναι μια απλή τροχιά αναφοράς σε σχέση με το ζεύγος (q_i, q_{i+1}) . Στην Εικόνα 2-19(α) απεικονίζεται αυτή η έννοια. T_1 και T_2 είναι απλές τροχιές αναφοράς σε σχέση με q_i, q_{i+1} . T_3 δεν είναι τροχιά αναφοράς, διότι δεν πέφτει στο κύκλο με κέντρο q_{i+1} . T_4 δεν είναι απλή τροχιά αναφοράς, υπάρχει ένα σημείο $a \in T_4$ έξω από τη διαδρομή της τροχιάς αναφοράς. Έχοντας υπόψη ένα ζευγάρι (q_i, q_{i+1}) , μπορούμε να ψάξουμε για απλές τροχιές αναφοράς με τη βοήθεια του R-tree index που έχει κατασκευαστεί στην προεπεξεργασία.

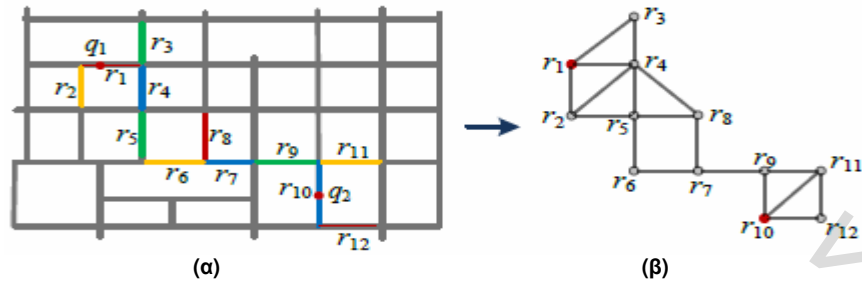


Εικόνα 2-19: Τροχιά αναφοράς

Spliced Reference Trajectory (Συνδεδεμένη Τροχιά Αναφοράς): Σε μια περιοχή με αραιά ιστορικά στοιχεία, η απαίτηση για μια τροχιά να προκριθεί μια απλή αναφορά είναι πολύ αυστηρή, γεγονός που καθιστά τον αριθμό λαμβανόμενων τροχιών αναφοράς πολύ μικρό για να στηρίξει το συμπέρασμα. Η έννοια της αναφοράς περιλαμβάνει ένα άλλο είδος τροχιών. Εξετάζοντας την Εικόνα 2-19(β) όπου ούτε T_1 ούτε T_2 , είναι μια απλή αναφορά. Αλλά όταν τα συνδέω, μια εικονική τροχιά T_3 θα σχηματιστεί, η οποία στην πραγματικότητα είναι μια καλή αναφορά για την εκτίμηση της πραγματικής διαδρομής του ερωτήματος. T_3 στην Εικόνα 2-19(β) είναι μια συνδεδεμένη τροχιά αναφοράς. Ουσιαστικά, μια συνδεδεμένη τροχιά αναφοράς είναι ένας συνδυασμός δύο τροχιών T_1 (από q_i) και T_2 (προς q_{i+1}) το οποίο έχει κάποια επικάλυψη. Όπως μπορούμε να δούμε από την Εικόνα 2-19(β), T_4 θα πρέπει να θεωρηθεί ως σημείο αναφοράς, αν επεκτείνουμε την ακτίνα από ϕ_1 σε ϕ_2 , αλλά δεν είναι λογικό να εκτιμηθεί η διαδρομή από q_i στο q_{i+1} καθώς είναι πολύ μακριά. Η αναζήτηση για συνδεδεμένες τροχιές αναφοράς μπορεί να γίνει ως εξής. Πρώτον, παίρνουμε δύο σύνολα υποψήφιων τροχιών εκδίδοντας δύο ερωτήματα και αφαιρώντας τις απλές τροχιές αναφοράς. Στη συνέχεια, θα βρούμε όλα τα συνδεδεμένα ζεύγη με την εκτέλεση χωρικής ένωσης στα δύο υποψήφια σύνολα.

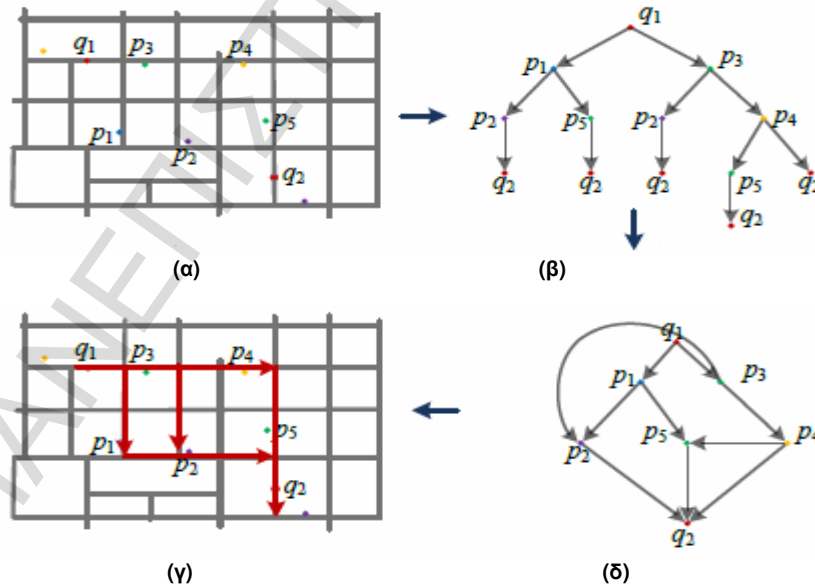
Συμπέρασμα Τοπικής Διαδρομής (Local Route Inference). Προτείνονται δύο προσεγγίσεις. Η *traverse graph based* προσέγγιση κατασκευάζει ένα εννοιολογικό διάγραμμα, χρησιμοποιώντας τμήματα διαδρομής που έχει διανύσει από ορισμένες τροχιές αναφοράς. Η *nearest neighbor based* προσέγγιση υιοθετεί ένα ευρετικό τρόπο να μεταφέρει από ένα σημείο αναφοράς στους κοντινότερους γείτονες μέχρι τον προορισμό. Μια υβριδική προσέγγιση που συνδυάζει τα πλεονεκτήματα των δύο μεθόδων για να επιτευχθεί καλύτερη αποτελεσματικότητα [ZZXZ2012].

Traverse Graph based Approach (Προσέγγιση βασισμένη στην Διάσχιση Γράφου): Βασική υπόθεση είναι ότι τα ιστορικά δεδομένα μπορεί να αντανakλούν τη δημοτικότητα διαδρομών. Αν ένα τμήμα του δρόμου δεν είναι ταξιδεμένη από καμία αναφορά, υπάρχει πιθανότητα το ερώτημα να μην περνά από αυτό. Δίνονται δύο διαδρομές R_a , R_b μεταξύ ορισμένων περιοχών, αν R_a είναι η συντομότερη διαδρομή, αλλά δεν ταξίδεψε καμία αναφορά, ενώ R_b διασχίζεται σε μεγάλο βαθμό, αλλά μακρύτερα από R_a , έχουμε περισσότερη σιγουριά ότι το R_b ταξίδεψε από το ερώτημα. Σε σύγκριση με το αρχικό οδικό δίκτυο, το γράφημα είναι πιο σύντομο, διότι αποτελείται από ακμές που έχουν διέλθει από κάποια αναφορά. Ο αλγόριθμος θα σαρώσει όλες τις τροχιές αναφοράς μία φορά για να πάρει το σετ ακμών διάσχισης (TE). Χτίζεται το γράφημα διάσχισης βρίσκοντας την λ -γειτονιά για κάθε ακμή $r \in TE$. Θέτει τις υποψήφιες ακμές (q_i, q_{i+1}) ως πηγές και προορισμούς και βρίσκει τις top-K συντομότερες διαδρομές. Τέλος σχεδιάζει αυτές τις διαδρομές για το αρχικό δίκτυο και παίρνει τις φυσικές διαδρομές. Η Εικόνα 2-20 απεικονίζει το γράφημα διάσχισης. Υπάρχουν 4 τοπικές τροχιές αναφοράς μεταξύ q_i και q_{i+1} . Στις ακμές διάσχισης δίνονται διαφορετικά χρώματα για να διακρίνονται οι τροχιές αναφοράς. Θέτοντας $\lambda=2$, κάθε ακμή διάσχισης συνδέεται με αυτά εντός ενός βήματος. Με την απομάκρυνση περιττών ακμών το γράφημα διάσχισης κατασκευάζεται όπως φαίνεται στην Εικόνα 2-20(b).



Εικόνα 2-20: Κατασκευή γραφήματος διάσχισης

Nearest Neighbor based Approach (Προσέγγιση βασισμένη στον Πλησιέστερο Γείτονα): Στη TGI προσέγγιση επιλέγοντας κατάλληλο λ είναι σημαντικό. Αν λ είναι μικρό η συνδεσιμότητα του γραφήματος είναι χαμηλή. Ένα μεγάλο λ κάνει τον αλγόριθμο αναποτελεσματικό, καθώς η αναζήτηση για την λ -γειτονιά για κάθε ακμή είναι δαπανηρή. Το ζήτημα είναι σοβαρό όταν η μέση απόσταση από τα σημεία αναφοράς είναι μεγάλη καθώς αποτελεσματικότητα και αποδοτικότητα θα χαθούν. Για να αντιμετωπιστεί, υπάρχει ένας άλλος αλγόριθμος, ο NNI, που είναι πιο προσαρμοστικός με την απόσταση σημείων αναφοράς. Η βασική ιδέα είναι, να βρούμε έναν τρόπο για να ταξιδέψει από q_i στο q_{i+1} μέσω σημείων αναφοράς, ξεκινώντας από q_i και αναζητώντας για το πλησιέστερο γείτονα που θα οδηγήσει στο q_{i+1} . Επαναλαμβάνει αυτή τη διαδικασία μέχρι ο προορισμός q_{i+1} να επιτευχθεί. Η προσέγγιση μπορεί να γίνει χρονοβόρα, όταν ο αριθμός αναφερόμενων σημείων γίνεται μεγάλος. Στην Εικόνα 2-21(α), τα σημεία με διαφορετικά χρώματα ανήκουν σε διαφορετικές τροχιές αναφοράς. Παρακολουθώντας τις επαναλήψεις μπορούμε να πάρουμε ένα δέντρο, Εικόνα 2-21(β). Συνολικά πρέπει να εκτελέσει 8 αναζητήσεις. Μερικά μπορεί να αποθηκευτούν αφού υπάρχουν κοινές υποδομές, $p_2 \rightarrow q_2$ και $p_5 \rightarrow q_2$. Μπορούμε να μοιραστούμε τις κοινές υποδομές, κρατώντας τις προηγούμενες διακλαδώσεις. Αν το επόμενο σημείο που εξετάσουμε είναι ήδη ένας κόμβος στο δέντρο, συνδέστε το τρέχον σημείο στον κόμβο. Δημιουργούμε ένα γράφημα διέλευσης που δείχνει τους τρόπους για να μεταβιβαστεί από q_1 στο q_2 . Όπως φαίνεται στην Εικόνα 2-21(δ), 6 αναζητήσεις απαιτούνται μετά την ανταλλαγή κοινών υποδομών. Έχουμε απαριθμήσει τις πιθανές διαδρομές από q_1 στο q_2 και συμπεράνει τη διαδρομή με τα σημεία αναφοράς, Εικόνα 2-21(γ).



Εικόνα 2-21: Nearest neighbour based inference

Hybrid Approach (Υβριδική Προσέγγιση): Η TGI προσέγγιση χρησιμοποιεί μια σταθερή ακτίνα για να αναζητήσει διασχισμένες γειτονικές ακμές. Έτσι, θα γίνει λιγότερο αποτελεσματική όταν η πυκνότητα των σημείων αναφοράς είναι χαμηλή. Από την άλλη, η NNI βρίσκει τους k

πλησιέστερους γείτονες σε κάθε σημείο, οπότε, δεν είναι ευαίσθητος στην κοινή απόσταση των σημείων αναφοράς. Όταν η πυκνότητα είναι υψηλή, ο NNI θα είναι χρονοβόρος λόγω της επαναληπτικής αναζήτησης. Για να είναι το σύστημα πιο προσαρμοστικό στην ετερογενή κατανομή των τροχιών σε πραγματική εφαρμογή, προτείνεται μια υβριδική προσέγγιση, η οποία χρησιμοποιεί ένα όριο τ σε οποία από τις δύο προσεγγίσεις θα υιοθετηθεί. Πριν εκτελώντας την τοπική διαδρομή συμπεράσματος, εκτιμάτε η πυκνότητα των σημείων αναφοράς από την αναλογία μεταξύ του αριθμού των σημείων και της περιοχής του Ελάχιστου Ορίου των σημείων. Αν η πυκνότητα είναι χαμηλότερη από τ , ο TGI θα επιλεγεί, αλλιώς θα επιλεγεί ο NNI.

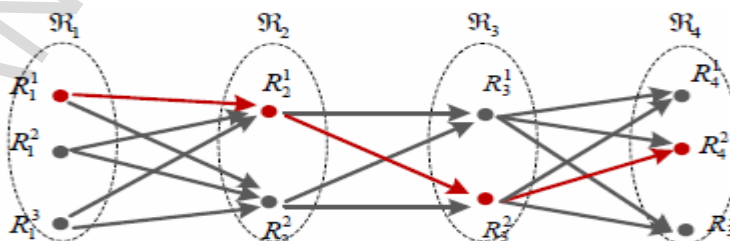
Συμπέρασμα Καθολικής Διαδρομής (Global Route Inference). Η έξοδος της τελευταίας φάσης είναι ένα σύνολο από τοπικές διαδρομές για κάθε ζεύγος (q_i, q_{i+1}) . Θα πρέπει να συνδεόνται διαδοχικές τοπικές διαδρομές για την απόκτηση της καθολικής διαδρομής. Δεδομένου ότι ο αριθμός των καθολικών διαδρομών μπορεί να είναι τεράστιος, είναι άχρηστο να προτείνονται όλα αυτά τα αποτελέσματα, η μέτρηση της ποιότητας διαδρομής είναι επιθυμητή, έτσι ώστε το τελικό αποτέλεσμα να μπορεί να διακριθεί. Προτείνεται η συνάρτηση βαθμολόγησης για τις καθολικές διαδρομές ενσωματώνοντας την δημοτικότητα τοπικών διαδρομών και ο αλγόριθμος δυναμικού προγραμματισμού για να υπολογίσει τις top-K καθολικές διαδρομές [ZZXZ2012].

Route Scoring Function (Συνάρτηση Βαθμολόγησης Διαδρομής): Μια καθολική διαδρομή αποτελείται από μια ακολουθία τοπικών διαδρομών, έτσι η ποιότητα του πρέπει να εξετάσει δύο πτυχές: 1) η ποιότητα της κάθε τοπικής διαδρομής 2) η ποιότητα των συνδέσεων μεταξύ των διαδοχικών τοπικών διαδρομών. Αντίστοιχα, προτείνονται η συνάρτηση δημοτικότητας για κάθε τοπική διαδρομή και η συνάρτηση μετάβασης εμπιστοσύνης για τη σύνδεσή τους. Η ποιότητα της τοπικής διαδρομής χαρακτηρίζεται από μια συνάρτηση δημοτικότητας η οποία είναι ένας δείκτης του πόσο συχνά η διαδρομή διανύεται από τις τροχιές αναφοράς. Προφανώς, όσο περισσότερες τροχιές αναφοράς διανύουν την διαδρομή, τόσο μεγαλύτερη δημοτικότητα έχει.

Top-K Global Routes (Top-K Καθολικές Διαδρομές): Δίνεται μια σειρά από σύνολα τοπικών διαδρομών (R_1, R_2, \dots, R_n) , βρίσκοντας τις top-K καθολικές διαδρομές όσον αφορά τη βαθμολογία τους. Υποθέτοντας ότι ο μέσος αριθμός στοιχείων κάθε συνόλου τοπικής διαδρομής είναι m , το μήκος της τροχιάς ερωτήματος είναι n , ο αριθμός συνδυασμού για την καθολική διαδρομή θα είναι m^{n-1} . Προτείνεται ο αλγόριθμος δυναμικού προγραμματισμού, K-GRI (top-K Global Route Inference) για να βρει τις top-K καθολικές διαδρομές αποτελεσματικά. Η βασική δομή της μεθόδου είναι ένα matrix $M[i, j]$. Η καταχώρηση $M[i, j]$ διατηρεί K καθολικές διαδρομές με την υψηλότερη βαθμολογία μεταξύ εκείνων που τελειώνουν με R_i^j , δηλ., $(* \diamond R_i^j) = \{R_1 \diamond R_2 \diamond \dots \diamond R_i^j, \forall R_m \in R_m, 1 \leq m < i\}$. Η καταχώρηση $M[i+1, k]$ μπορεί να προέρχεται από τον ακόλουθο τύπο.

$$M[i+1, k] = \text{top}K_{j=1}^{|\mathcal{R}_i|} \{R | R = R' \diamond R_{i+1}^k, \forall R' \in M[i, j]\}$$

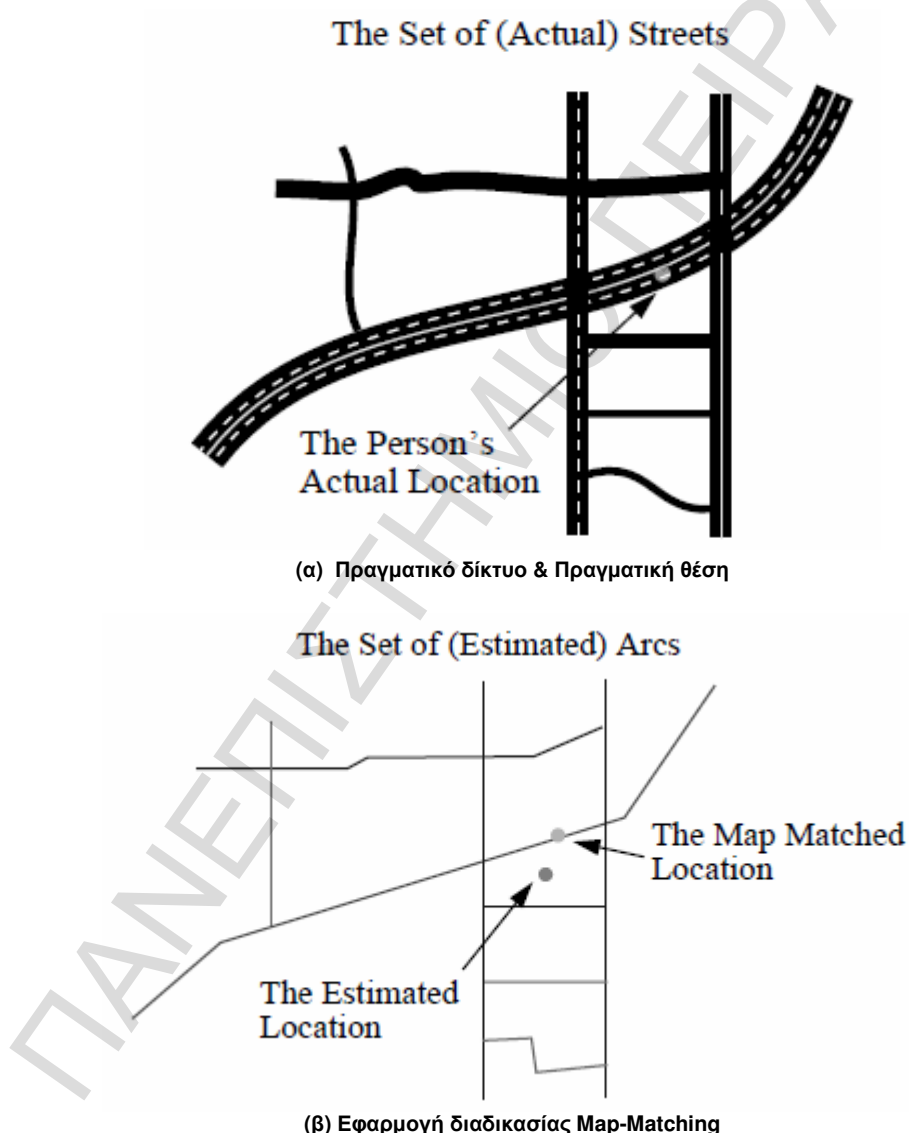
Στην Εικόνα 2-22, εάν η συνολική διαδρομή $R_4 = R_1^1 \diamond R_2^1 \diamond R_3^2 \diamond R_4^2$ είναι με υψηλότερη βαθμολογία, τότε είναι σίγουρο ότι $R_3 = R_1^1 \diamond R_2^1 \diamond R_3^2$ είναι επίσης top-1 μεταξύ $* \diamond R_3^2$. Αλλιώς, μπορούμε να αντικαταστήσουμε το R_3 από εκείνη της οποίας η βαθμολογία είναι υψηλότερη, συμβολίζεται με R_3 για να παίρνεται μια καλύτερη $R_4 = R_3 \diamond R_4^2$ διαδρομή [ZZXZ2012].



Εικόνα 2-22: Global route inference

2.10 Map Matching

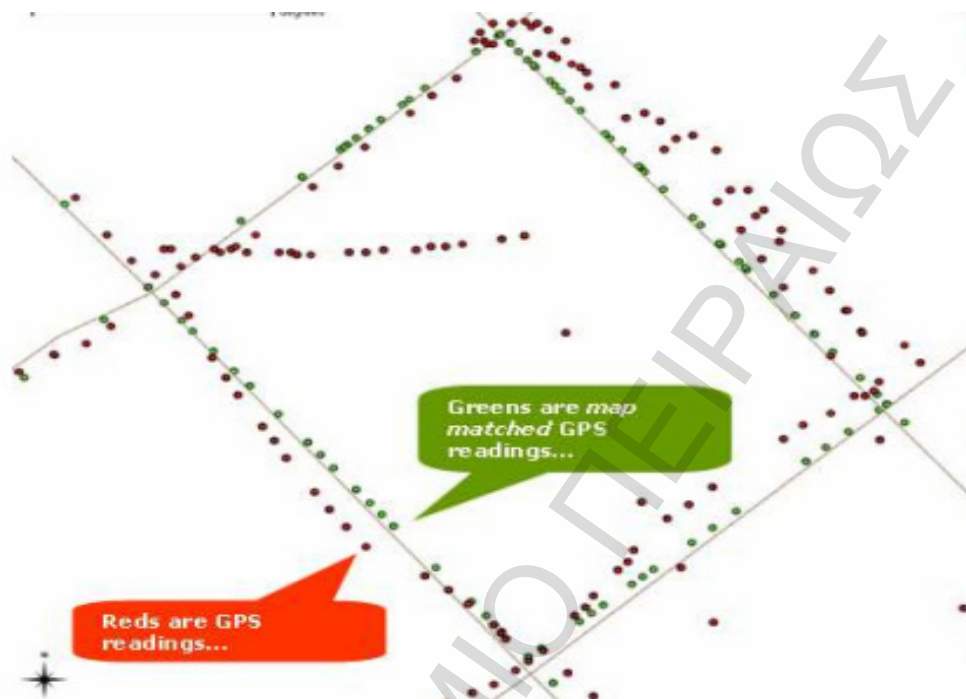
Το map-matching είναι μια διαδικασία που έχει στόχο να προσεγγίζει όσο πιο πιστά γίνεται την τροχιά ενός κινούμενου αντικείμενου πάνω σε έναν χάρτη. Θεωρούμε ότι ένα αντικείμενο κινείται μέσα σε ένα πεπερασμένο σύνολο διαδρομών και για ένα πεπερασμένο σύνολο χρονικών στιγμών μπορούμε να γνωρίζουμε τη θέση του αντικείμενου. Στόχος είναι να καθοριστεί η θέση του αντικείμενου τη χρονική στιγμή t . Όπως βλέπουμε στην Εικόνα 2-23 (α) είναι το πραγματικό δίκτυο και μια πραγματική θέση πάνω σε αυτό. Ενώ στην Εικόνα 2-23 (β) την map-matched θέση και το δίκτυο σε μια γραφική αναπαράσταση που αποτελείται από καμπύλες. Κάθε καμπύλη μπορεί να χαρακτηριστεί από μια πεπερασμένη ακολουθία σημείων τα οποία ανήκουν στο δισδιάστατο χώρο. Τα σημεία αυτά λέγονται κόμβοι και είναι σημεία αφετηρίας ή τερματισμού ή διασταύρωσης καμπυλών. Το πρόβλημα αυτό ονομάζεται map-matching επειδή στόχος είναι η προσπάθεια να «ταιριάξουμε» τη εκτιμωμένη θέση με μια καμπύλη στο χάρτη και να εκτιμηθεί με μεγαλύτερη ακρίβεια η πραγματική θέση του χρήστη [BK1996].



Εικόνα 2-23: Το Map-Matching Πρόβλημα

Στην Εικόνα 2-24 βλέπουμε ένα στιγμιότυπο της λειτουργίας της διαδικασίας του map-matching όπου είναι εύκολο να διακρίνουμε πως αξιοποιούμε τις τεχνικές του map-matching. Βλέποντας την παρακάτω εικόνα μπορούμε να διακρίνουμε ότι οι εκτιμήσεις του GPS (τα

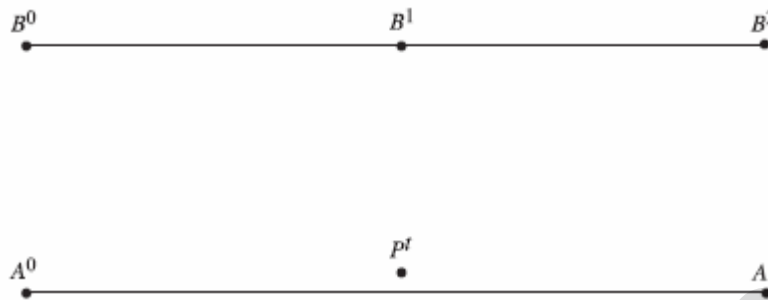
κόκκινα σημεία) είναι εκτός του οδικού δικτύου, υποθέτοντας ότι οι γραμμές είναι οι δρόμοι του οδικού δικτύου. Με την βοήθεια της διαδικασίας map-matching μεταφέρουμε τις εκτιμήσεις στο οδικό δίκτυο και ως αποτέλεσμα έχουμε μια καλύτερη εικόνα της τροχιάς που ακολουθήθηκε (τα πράσινα σημεία). Η διαδικασία αυτή προσφέρει μεγαλύτερη ακρίβεια για την πραγματική τροχιά που έχει ακολουθηθεί.



Εικόνα 2-24: Το Map-Matching

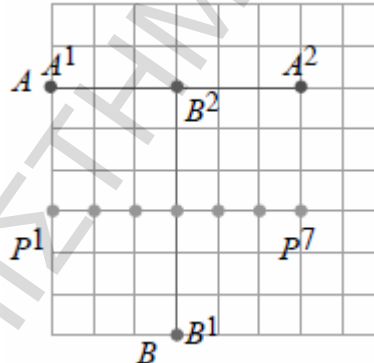
Υπάρχει ένας αριθμός από διαφορετικούς τρόπους για να προσεγγίσει το map matching πρόβλημα, καθένας από τους οποίους έχει πλεονεκτήματα και μειονεκτήματα [WBK2000].

Map matching ως πρόβλημα αναζήτησης. Κάποιος μπορεί να δει το map matching πρόβλημα ως ένα απλό πρόβλημα αναζήτησης. Τότε το πρόβλημα είναι να ταιριάζει το P^t με το «πλησιέστερο κόμβο» ή σημείο στο δίκτυο. Ένας αριθμός δομών δεδομένων και αλγορίθμων υπάρχουν για τον εντοπισμό όλων των σημείων κοντά σε ένα δεδομένο σημείο (που συχνά αποκαλείται ένα ερώτημα απόστασης). Είναι στη συνέχεια ένα απλό θέμα για να βρουν την απόσταση μεταξύ P^t και κάθε κόμβου και σημείου που είναι μέσα σε μια “λογική” απόσταση από αυτό (ανεξάρτητα από την χρησιμοποιούμενη μονάδα μέτρησης), και επιλέγει το πιο κοντινό. Ενώ αυτή η προσέγγιση είναι λογικά τόσο εύκολη να εφαρμοστεί και τόσο γρήγορη, έχει πολλά προβλήματα στην πράξη. Ίσως το πιο σημαντικό, εξαρτάται κριτικά σχετικά με τον τρόπο με τον οποίο τα σημεία σχήματος χρησιμοποιούνται στο δίκτυο. Βλέποντας το παράδειγμα που φαίνεται στην Εικόνα 2-25. Το P^t είναι πολύ πιο κοντά στο B^1 από τι είναι είτε στο A^0 είτε στο A^1 , ως εκ τούτου, θα πρέπει να ταιριάζει με το τόξο B, ακόμη και αν είναι διαισθητικά σαφές ότι θα πρέπει να ταιριάζει με το τόξο A. Ως εκ τούτου, αυτό το είδος του αλγορίθμου είναι πολύ ευαίσθητο στον τρόπο με τον οποίο το δίκτυο είχε ψηφιοποιηθεί. Δηλαδή, τα τόξα με περισσότερα σημεία σχήματος είναι πιο πιθανό να ταιριάζουν. Θα μπορούσε κανείς να υποστηρίξει ότι αυτό το πρόβλημα θα μπορούσε να ξεπεραστεί απλά με τη συμπεριληψη περισσότερων σημείων σχήματος για κάθε τόξο. Δυστυχώς, αυτό αυξάνει δραματικά το μέγεθος του δικτύου και δεν είναι εγγυημένη για να διορθώσει το πρόβλημα [WBK2000].



Εικόνα 2-25: Ένα πρόβλημα με point-to-point matching

Map matching ως στατιστική εκτίμηση. Κάποιος μπορεί επίσης να δει το map matching ως πρόβλημα στατιστικής εκτίμησης. Στην προσέγγιση αυτή, κάποιος θεωρεί μια ακολουθία σημείων $P^0 \dots P^i$ και επιχειρεί να φτιάξει μια καμπύλη με αυτά. Αυτή η καμπύλη περιορίζεται να κινείται στο δίκτυο. Τέτοιου είδους προσέγγιση έχει διερευνηθεί σε πολυάριθμες εργασίες και είναι αρκετά ελκυστική. Είναι ιδιαίτερα κομψό όταν το μοντέλο που περιγράφει την "φυσική της κίνησης" είναι απλό (π.χ., η κίνηση είναι δυνατή μόνο κατά μήκος μιας ευθείας γραμμής). Δυστυχώς, στις περισσότερες πρακτικές εφαρμογές, η φυσική της κίνησης υπαγορεύεται από (ή να περιορίζεται από) το δίκτυο. Αυτό το καθιστά αρκετά δύσκολο το μοντέλο. Για να καταλάβουμε γιατί αυτό είναι σημαντικό, εξετάζεται το δίκτυο που φαίνεται στην Εικόνα 2-26. Σε αυτό το παράδειγμα, οι θέσεις $P^1 \dots P^7$ έχουν καταγραφεί. Στόχος είναι να σχηματιστεί μια καμπύλη σε αυτά τα σημεία, αλλά η καμπύλη περιορίζεται να κινείται στο δίκτυο. Σε αυτή την περίπτωση, υπάρχουν δύο καμπύλες υποψήφιας, A και B (αγνοούμε το υπόλοιπο του δικτύου για λόγους απλότητας). Σε γενικές γραμμές, η καμπύλη P είναι πλησιέστερα στην καμπύλη B από ό, τι είναι στην καμπύλη A. Έτσι, αν κάποιος χρησιμοποιεί ένα απλό μοντέλο της κίνησης θα ταιριάζει το P στην B και όχι στην A [WBK2000].

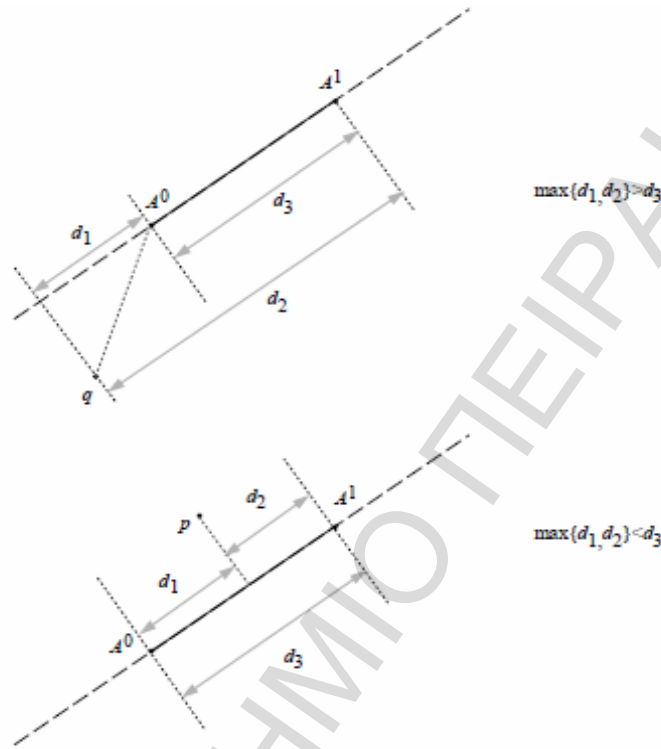


Εικόνα 2-26: Εφαρμογή καμπύλης

Ο συνδυασμός της απλότητας της απλής προσέγγισης αναζήτησης με ορισμένες από τις ιδέες της στατιστικής προσέγγισης θα εφαρμοστεί για τέσσερις διαφορετικούς αλγορίθμους [WBK2000].

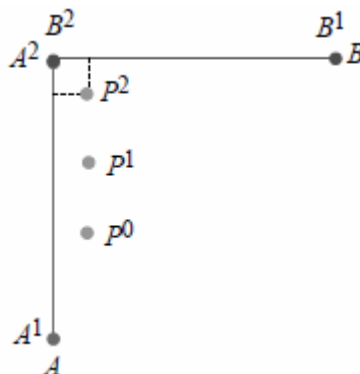
Αλγόριθμος 1 είναι πολύ απλός. Βρίσκει τους κόμβους που είναι κοντά στο GPS "tick" και βρίσκει το σύνολο των τόξων που προσπίπτουν σε αυτούς τους κόμβους. Βρίσκει τα πιο κοντινά από αυτά τα τόξα και σχεδιάζει το σημείο σε αυτό το τόξο (χρησιμοποιώντας μια ελάχιστη προβολή κανόνα). Όπως δείχνεται στην Εικόνα 2-27, ο υπολογισμός της ελάχιστης απόστασης μεταξύ ενός σημείου και ενός τμήματος της γραμμής είναι ελαφρώς πιο περίπλοκος από τον υπολογισμό της ελάχιστης απόστασης μεταξύ ενός σημείου και μιας γραμμής. Ο υπολογισμός της ελάχιστης απόστασης μεταξύ p και του τμήματος της γραμμής A^0 και A^1 είναι απλός, δεδομένου ότι είναι ίδια με την ελάχιστη απόσταση μεταξύ p και της γραμμής A^0 και A^1 . Ωστόσο, όταν υπολογίζουμε την απόσταση μεταξύ q και της γραμμής A^0 και A^1 , βλέπουμε ότι η "κάθετη" τέμνει τη γραμμή έξω από το τμήμα της γραμμής. Ως εκ τούτου, θα πρέπει επίσης να υπολογίσει την απόσταση μεταξύ q και A^0 και A^1 και να επιλέξει την μικρότερη. Τέλος, δεδομένου ότι κάθε

τόξο είναι μια τμηματικά γραμμική καμπύλη, πρέπει να βρούμε την ελάχιστη απόσταση από το σημείο του ενδιαφέροντος για κάθε ένα από τα ευθύγραμμα τμήματα που περιέχουν το A και επιλέγεται το μικρότερο. Έτσι, ο υπολογισμός της ελάχιστης απόστασης ανάμεσα σε ένα σημείο, P^i , και ένα τόξο A, προϋποθέτει την εξεύρεση της ελάχιστης απόστασης μεταξύ P^i και των τμημάτων της γραμμής $\{\lambda A^0 + (1-\lambda)A^1, \lambda \in [0,1]\}$, $\{\lambda A^1 + (1-\lambda)A^2, \lambda \in [0,1]\}$, ... , $\{\lambda A^{n-1} + (1-\lambda)A^n, \lambda \in [0,1]\}$ και επιλέγοντας το μικρότερο [WBK2000].



Εικόνα 2-27: Η απόσταση μεταξύ ενός σημείου και ενός τμήματος

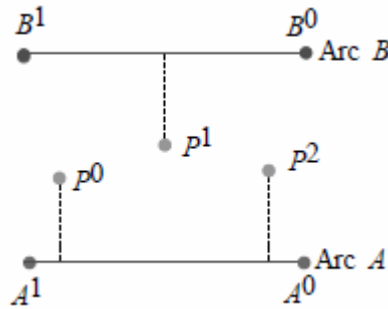
Ο αλγόριθμος αυτός έχει πολλές ελλείψεις. Πρώτον, δεν κάνει χρήση “ιστορικών” πληροφοριών και αυτό μπορεί να προκαλέσει προβλήματα του είδους που απεικονίζονται στην Εικόνα 2-28. Η εκτιμώμενη θέση P^2 είναι εξίσου κοντά στα τόξα A και B. Ωστόσο, δεδομένου των P^0 και P^1 είναι σαφές ότι το P^2 πρέπει να ταιριάζει με το τόξο A.



Εικόνα 2-28: Ένα πρόβλημα αντιστοίχισης σημείου-προς-καμπύλη

Ένα άλλο πρόβλημα με αυτό τον αλγόριθμο είναι ότι μπορεί να είναι αρκετά “ασταθής”. Αυτό απεικονίζεται στην Εικόνα 2-29. Τα σημεία P^0 , P^1 , και P^2 είναι όλα σε ίση απόσταση από τα τόξα A και B. Ωστόσο, αποδεικνύεται ότι P^0 και P^2 είναι ελαφρώς πιο κοντά στο A και P^1 είναι

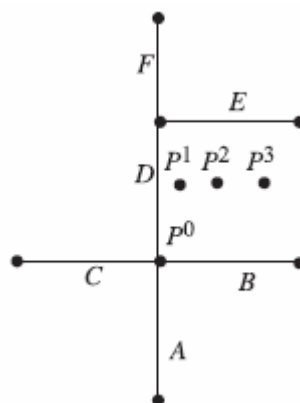
ελαφρώς πιο κοντά στο B. Ως εκ τούτου, η αντιστοίχιση (matching) ταλαντεύεται μπρος και πίσω μεταξύ των δύο. Ο αλγόριθμος 1 θα παίξει το ρόλο ενός "straw man". Ο αλγόριθμος 1 είναι γρήγορος και εύκολος στην εφαρμογή.



Εικόνα 2-29: Ένα άλλο πρόβλημα αντιστοίχισης σημείου-προς-καμπύλη

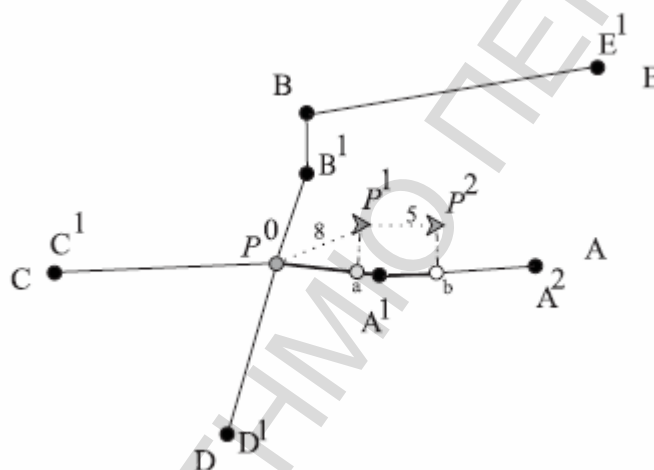
Αλγόριθμος 2 είναι ταυτόσημος με τον αλγόριθμο 1 εκτός του ότι θα κάνει χρήση της πληροφορίας "heading". Εάν η επικεφαλίδα δεν είναι συγκρίσιμη με την επικεφαλίδα του τόξου, τότε, το τόξο απορρίπτεται. Έτσι, για παράδειγμα, ένα GPS tick δεν θα ταιριάζει με ένα τόξο που είναι κάθετο στην τρέχουσα κατεύθυνση κίνησης (όπως συνέβη στο παράδειγμα στην εικόνα 2-29). Και πάλι, ο αλγόριθμος αυτός περιλαμβάνεται πρωτίστως ως ένας "straw man" [WBK2000].

Αλγόριθμος 3 είναι μια παραλλαγή του αλγόριθμου 2 που χρησιμοποιεί τοπολογικές πληροφορίες. Συγκεκριμένα, ενώ οι Αλγόριθμοι 1 και 2 χρησιμοποιούν μόνο ένα ερώτημα απόστασης για να εντοπίσουν υποψήφιους κόμβους (και, ως εκ τούτου, υποψήφια τόξα), ο Αλγόριθμος 3 χρησιμοποιεί επίσης πληροφορίες σύνδεσης. Ειδικά, αν ο αλγόριθμος έχει εμπιστοσύνη στο προηγούμενο matching, θα χρησιμοποιήσει την τοπολογία του δικτύου για τον εντοπισμό υποψήφιων κόμβων για το επόμενο matching. Από την άλλη πλευρά, εάν ο αλγόριθμος δεν έχει εμπιστοσύνη στο προηγούμενο matching, θα χρησιμοποιήσει ένα ερώτημα απόστασης. Ένας τρόπος για να κατανοήσουμε πώς τοπολογικές πληροφορίες μπορούν να χρησιμοποιηθούν είναι το παράδειγμα που φαίνεται στην Εικόνα 2-30. Ας υποθέσουμε ότι γνωρίζουμε ότι το άτομο ήταν αρχικά στο P^0 . Τότε γνωρίζουμε ότι το P^1 μπορεί να είναι μόνο στην A, B, C, ή D. Στην πραγματικότητα, δίνεται ένα επαρκώς μικρό χρονικό διάστημα μεταξύ των μετρήσεων, θα μπορούσε επίσης να γνωρίζουν ότι P^3 μπορεί μόνο να συνδυαστεί με A, B, C, ή D. Αυτό το είδος των πληροφοριών θα μπορούσε να αποτρέψει από το λάθος να ταιριάζει το P^3 με, ας πούμε E. Αυτός είναι ευκολότερος τρόπος για να ενσωματώσει τη φυσική της κίνησης σε έναν αλγόριθμο που βασίζεται στην αναζήτηση [WBK2000].



Εικόνα 2-30: Χρησιμοποιώντας τοπολογικές πληροφορίες

Αλγόριθμος 4 χρησιμοποιεί καμπύλη - προς - καμπύλη matching. Πρώτον, εντοπίζει υποψήφιους κόμβους χρησιμοποιώντας τις ίδιες τεχνικές όπως ο Αλγόριθμος 3. Στη συνέχεια, δίνεται έναν υποψήφιο κόμβο, κατασκευάζει τμηματικά γραμμικές καμπύλες από το σύνολο των διαδρομών που προήλθε από αυτόν τον κόμβο. Δεύτερον, κατασκευάζει μια τμηματικά γραμμική καμπύλη χρησιμοποιώντας τα σημεία (P^s, \dots, P^l) και υπολογίζει την απόσταση μεταξύ αυτής της καμπύλης και τις καμπύλες που αντιστοιχούν στο δίκτυο. Τέλος, επιλέγει την πλησιέστερη καμπύλη και σχεδιάζει το σημείο επάνω σε εκείνη την καμπύλη. Οι λεπτομέρειες αυτού του αλγορίθμου απεικονίζονται στην Εικόνα 2-31. Τα P^1 και P^2 , αντιπροσωπεύουν τα προηγούμενα και τρέχοντα σημεία, αντίστοιχα. Το σημείο P^0 είναι ένας υποψήφιος κόμβος. Η διακεκομμένη γραμμή είναι η τμηματικά γραμμική καμπύλη, P , κατασκευασμένο από P^0, P^1 και P^2 . Τα δύο τμήματα της παρούσας τμηματικά γραμμικής καμπύλης είναι 8 και 5 μονάδες σε μήκος, αντίστοιχα. Υπάρχουν πολλοί τρόποι για να υπολογίσει την απόσταση μεταξύ των δύο καμπυλών. Συγκεκριμένα, στη Εικόνα 2-31, για να υπολογίσουν την απόσταση μεταξύ της καμπύλης και του A χρησιμοποιούνται τα δύο σημεία a και b . Το a είναι το σημείο που είναι 8 μονάδες κατά μήκος του τόξου A ξεκινώντας από P^0 και b είναι το σημείο που είναι 5 μονάδες από το a κατά μήκος του τόξου A . Η απόσταση μεταξύ P και A στη συνέχεια υπολογίζεται ως $\|P^1 - a\|_2 + \|P^2 - b\|_2$. Μία παρόμοια διαδικασία χρησιμοποιείται για όλα τα C και D , και για άλλους υποψήφιους κόμβους. Το P^2 προβάλλεται πάνω στο πιο κοντινό [WBK2000].



Εικόνα 2-31: Αποστάσεις Καμπύλη-προς-καμπύλη

Οι τύποι δεδομένων που χειρίζεται το map matching περιλαμβάνουν point-to-line, line-to-line και polyline-to-polyline. Με βάση τα διαχρονικά χαρακτηριστικά, οι αλγόριθμοι map-matching μπορούν επίσης να ταξινομηθούν σε on-line map-matching και off-line map-matching [ZG2006]. Οι on-line map-matching μέθοδοι βασίζονται σε δεδομένα πραγματικού χρόνου (real-time). Οι off-line map-matching μέθοδοι βασίζονται στο σύνολο των δεδομένων που συλλέγονται. Στη εργασία [ZG2006] εξετάζονται οι map-matching μέθοδοι στο πλαίσιο των μελετών ταξιδιού / μεταφοράς.

Location based services (LBSs) καλύπτουν ένα ευρύ φάσμα υπηρεσιών παρέχοντας σχετικές πληροφορίες τοποθεσίας. Η πραγματική θέση του χρήστη δεν μπορεί ακριβώς να εμφανίζεται στον ψηφιακό χάρτη. Αυτό είναι ένα από τα σημαντικά θέματα και διαδραματίζει ουσιαστικό ρόλο στις LBSs. Αυτό το φαινόμενο ονομάζεται ως το map-matching πρόβλημα. Η απόδοση του map-matching προφανώς θα επηρεάσει την ποιότητα των Location Based Services. Matching με χαμηλή ακρίβεια μπορεί να οδηγήσει σε λάθος τοποθεσία ή να προκαλέσει μεγάλη απώλεια για παράδειγμα στις εμπορικές ή βιομηχανικές δραστηριότητες. Ως εκ τούτου, τα ζητήματα σχετικά με τις αβεβαιότητες θέσης έχουν δημιουργήσει ένα σημείο της έρευνας και σχετικά ζητήματα που πρέπει να επιλυθούν στο LBS. Στην εργασία [AWB2011] μελετώνται τα παραπάνω ζητήματα που αφορούν τις map-matching αβεβαιότητες θέσης. Στην εργασία αυτή, συζητήθηκαν οι τεχνολογίες του map matching και οι τεχνικές για το πρόβλημα της χαμηλής ακρίβειας του map matching. Είναι απαραίτητη προϋπόθεση για το χρήστη να αποκτήσει ακριβείς πληροφορίες θέσης. Εάν υπάρχει χαμηλή ακρίβεια θέσης, η ποιότητα των

υπηρεσιών θέσης μειώνεται σημαντικά. Σε γενικές γραμμές, οι LBSs ταξινομούνται σε υπηρεσίες εντοπισμού θέσης και υπηρεσίες ενημέρωσης θέσης. Η υπηρεσία εντοπισμού θέσης εκφράζει την τρέχουσα θέση του χρήστη σε ένα χάρτη. Το σύστημα υπηρεσιών ενημέρωσης θέσης προσφέρει χρήσιμες πληροφορίες για τον χρήστη ανάλογα με τη θέση του χρήστη. Για παράδειγμα, η αναζήτηση για το πλησιέστερο ξενοδοχείο. Δεν έχει σημασία ποια υπηρεσία είναι, η ακριβής θέση και το map matching απαιτούνται [AWB2011].

Στη συνέχεια θα γίνει ανάλυση για τους παράγοντες που επηρεάζουν το map-matching.

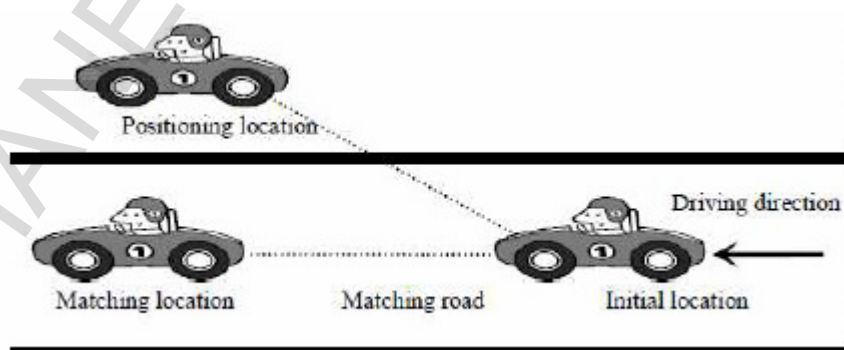
Μέθοδοι εντοπισμού θέσης (Positioning Methods). Ανάλογα με τις διαφορετικές ανάγκες των χρηστών, οι μέθοδοι εντοπισμού θέσης είναι διαφορετικοί. Υπάρχουν τρεις μέθοδοι εντοπισμού θέσης: point dynamic position, real time differential dynamic position και post processing differential position. Η μέθοδος point dynamic position καθορίζει το δέκτη σε ένα κινούμενο όχημα να αποκτήσει την θέση σε πραγματικό χρόνο και να δείξει την παρακολούθηση του οχήματος. Differential positioning θέτει σταθμό αναφοράς σε ένα ή περισσότερα σημεία που οι συντεταγμένες είναι γνωστές και την συνεχή παρακολούθηση.

Ψηφιακός Χάρτης (Digital Map). Στις location based services, ο ψηφιακός χάρτης είναι το θεμέλιο του όλου συστήματος. Μόνο ο ακριβής ψηφιακός χάρτης μπορεί να εμφανίσει τη σωστή διαδρομή πλοήγησης και τη θέση των ατόμων. Τα τελευταία χρόνια, με βελτιωμένες μεθόδους παραγωγής ψηφιακού χάρτη, η ακρίβεια βελτιώνεται σημαντικά. Ως αποτέλεσμα, η επίδραση του ψηφιακού χάρτη στο χάρτη αντιστοίχισης είναι λίγο.

Μετατροπή Συντεταγμένης (Coordinate Transformation). Το GPS χρησιμοποιεί το WGS84 σύστημα συντεταγμένων που είναι ένα γεωκεντρικό σύστημα συντεταγμένων. Ως εκ τούτου, οι μετρήσεις GPS πρέπει να μετατραπούν στο καρτεσιανό σύστημα συντεταγμένων.

Geocoding. Geocoding είναι η διαδικασία της ανάθεσης γεωγραφικών συντεταγμένων σε περιοχές ενδιαφέροντος το οποίο περιλαμβάνει δύο μέρη την διεύθυνση τυποποίησης και την διεύθυνση matching. Η Διεύθυνση Τυποποίησης προετοιμάζει τη διεύθυνση σε μια γνωστή μορφή διορθώνοντας τα ορθογραφικά λάθη και η διεύθυνση matching είναι να μετατρέψει την διεύθυνση τυποποίησης σε γεωγραφικές συντεταγμένες. Το Geocoding βελτιώνει την αναγνωσιμότητα των δεδομένων και την προβολή των δεδομένων στο χάρτη. Υπάρχουν δύο κύριες πηγές αβεβαιότητας στο geocoding, αυτή σχετικά με τη βάση δεδομένων που χρησιμοποιείται ως δεδομένα αναφοράς σε αντικείμενα geocoding και ένα σχετικά με τον αλγόριθμο διεύθυνσης matching. Το κλειδί της διεύθυνσης τυποποίησης είναι ο κατακερματισμός της διεύθυνσης. Στην πραγματικότητα, είναι να διαιρέσει τη διεύθυνση εισόδου συμβολοσειράς σε διαφορετικές φράσεις. Μετά την τυποποίηση διεύθυνσης, είναι απαραίτητο να κάνουμε την αντίστοιχη διεύθυνση.

Map matching. Το map matching είναι αυτό που ταιριάζει την τροχιά του οχήματος με την διαδρομή που παρέχεται από τον ψηφιακό χάρτη. Η αρχή του map matching φαίνεται στην Εικόνα 2-32. Η ακρίβεια των πληροφοριών για τον εντοπισμό της θέσης και ο ψηφιακός χάρτης είναι η βάση του map matching. Ο Map matching αλγόριθμος είναι ένα λογισμικό για τη βελτίωση του map matching.

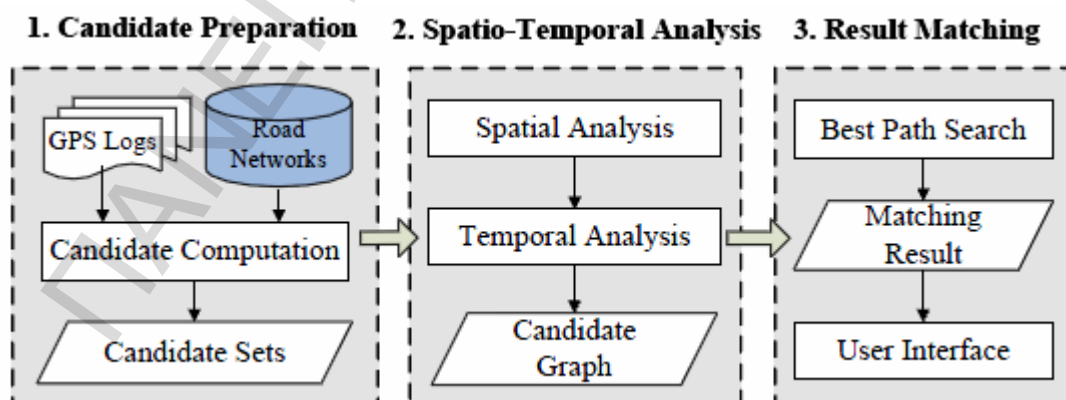


Εικόνα 2-32: Η αρχή του Map-Matching

Ο map matching αλγόριθμος επιτυγχάνει αντιστοίχιση μεταξύ του δείγματος και προτύπου. Το δείγμα είναι η τροχιά του οχήματος που προετοιμάζεται για να ταιριάζει; το πρότυπο είναι τα ακριβή δεδομένα δρόμο. Το Map-matching στηρίζεται σε δύο προϋποθέσεις: Πρώτα, ο ψηφιακός χάρτης είναι πιο ακριβής από την ακρίβεια της πλοήγησης. Δεύτερον, το κινητό τερματικό είναι τοποθετημένο επί της οδού. Ο map matching αλγόριθμος μπορεί να χωριστεί σε έναν map matching αλγόριθμο βασισμένο στην προβολή και σε έναν βασισμένο στην αναγνώριση προτύπων. Και οι δύο μέθοδοι έχουν τα πλεονεκτήματα και τα μειονεκτήματά τους. Ο αλγόριθμος βασισμένο στην προβολή είναι απλός και αποτελεσματικός, αλλά τα ποσοστά σφάλματος είναι σχετικά υψηλά σε μερικά περίπλοκα τμήματα του δρόμου. Ενώ, Ο αλγόριθμος βασισμένο στην αναγνώριση προτύπων είναι υψηλό ποσοστό, έχει το μειονέκτημα του πολύπλοκου υπολογισμού και φτωχό σε πραγματικό χρόνο.

Αξιολόγηση Αβεβαιοτήτων (Evaluation of Uncertainties). Γενικά οι αβεβαιότητες προέρχονται από δύο κύριες πηγές: φυσικά και τεχνητά σφάλματα. Οι παραπάνω παράγοντες συχνά αναφέρονται ως φυσικά σφάλματα. Τα τεχνητά σφάλματα έχουν να κάνουν με το πόσο καλά ο αλγόριθμος θα ταιριάζει τη θέση του χρήστη με το χάρτη [AWB2011].

Map-matching είναι η διαδικασία ευθυγράμμισης μιας ακολουθίας παρατηρούμενων θέσεων χρήστη με το οδικό δίκτυο σε έναν ψηφιακό χάρτη. Είναι ένα θεμελιώδες στάδιο προεπεξεργασίας για πολλές εφαρμογές, όπως η διαχείριση κινούμενου αντικείμενο, ανάλυση της ροής κυκλοφορίας. Στην πράξη, υφίσταται χαμηλός ρυθμός δειγματοληψίας GPS τροχιών (π.χ., ένα σημείο κάθε 2-5 λεπτά). Συνήθως, μια GPS τροχιά αποτελείται από μια ακολουθία σημείων με το γεωγραφικό πλάτος, γεωγραφικό μήκος, καθώς και πληροφορίες timestamp. Ωστόσο, αυτά τα δεδομένα δεν είναι ακριβής λόγω σφαλμάτων μέτρησης που προκαλείται από τον περιορισμό των συσκευών GPS και σφάλματος δειγματοληψίας που προκαλείται από το ρυθμό δειγματοληψίας. Συνεπώς, οι παρατηρούμενες GPS θέσεις συχνά χρειάζεται να ευθυγραμμιστούν με το οδικό δίκτυο σε ένα δεδομένο ψηφιακό χάρτη. Αυτή η διαδικασία ονομάζεται Map-matching. Οι υπάρχουσες map-matching προσεγγίσεις έχουν ασχοληθεί μόνο με υψηλό ρυθμό δειγματοληψίας GPS δεδομένων (συνήθως ένα σημείο κάθε 10-30s). Στην εργασία [LZZX2009], προτείνεται ένας νέος map-matching αλγόριθμος που ονομάζεται ST-Matching για χαμηλό ρυθμό δειγματοληψίας σε GPS τροχιές. Η εργασία αυτή ασχολείται με το πρόβλημα του σφάλματος δειγματοληψίας. Τα αποτελέσματα από τα πειράματα που διεξάχθηκαν τόσο σε συνθετικά όσο σε πραγματικά δεδομένα δείχνουν ότι ο ST-matching αλγόριθμος ξεπερνά σημαντικά άλλους αλγορίθμους όσον αφορά την ακρίβεια για χαμηλή δειγματοληψία σε τροχιές. Ενώ σε σύγκριση με το AFD-based αλγόριθμο, ο ST-Matching βελτιώνει επίσης την ακρίβεια και το χρόνο λειτουργίας. Αξιολογείται λόγω του χρόνου λειτουργίας και της ακρίβειας του map-matching. Η αρχιτεκτονική του προτεινόμενου map-matching συστήματος δείχνεται στην Εικόνα 2-33. Αποτελείται από τρία βασικά συστατικά: Candidate Preparation (Υποψήφια Προετοιμασία), Spatial and Temporal Analysis (Χωρική και Χρονική Ανάλυση), καθώς και Result Matching (Αποτελέσματα Matching) [LZZX2009].



Εικόνα 2-33: Επισκόπηση της αρχιτεκτονικής του συστήματος

Candidate Preparation: Αυτό το συστατικό περιέχει μια βάση δεδομένων με το οδικό δίκτυο. Αποδέχεται GPS τροχιά από τον χρήστη, και, στη συνέχεια ανακτά όλα τα πιθανά υποψήφια

σημεία για κάθε σημείο δειγματοληψίας στην τροχιά. Η έξοδος αυτού του συστατικού είναι ένα σύνολο υποψηφίων σημείων και των οδικών τμημάτων που βρίσκονται.

Spatial and Temporal Analysis: Αυτό το συστατικό εκτελεί την χωρική ανάλυση που ακολουθείται από τη χρονική ανάλυση των ανακτημένων υποψηφίων συνόλων και την τροχιά που πρέπει να συνδυαστεί. Η χωρική ανάλυση όχι μόνο εξετάζει την απόσταση μεταξύ ενός GPS σημείου και υποψήφια τμημάτων του δρόμου για αυτό το σημείο, αλλά λαμβάνει επίσης υπόψη τις τοπολογικές πληροφορίες του οδικού δικτύου. Για να αποφευχθούν τα κυκλικά μονοπάτια, χρησιμοποιείται η συντομότερη διαδρομή για να μετρηθεί η ομοιότητα μεταξύ κάθε υποψήφιας διαδρομής και της «πραγματικής» διαδρομής. Η χρονική ανάλυση μετρά την πραγματική μέση ταχύτητα μεταξύ οποιωνδήποτε γειτονικών σημείων. Συγκρίνει τη μέση ταχύτητα με τους περιορισμούς της τυπικής ταχύτητας σε κάθε υποψήφια διαδρομή. Οι πληροφορίες μπορούν να χρησιμοποιηθούν αργότερα για να ταιριάζει την τροχιά με την υποψήφια διαδρομή με τις περισσότερες παρόμοιες συνθήκες ταχύτητας κατά τη διάρκεια του χρονικού διαστήματος. Μετά την χωρική και χρονική ανάλυση, ένα υποψήφιο γράφημα κατασκευάζεται ως η έξοδος αυτού του συστατικού. Οι κόμβοι του γραφήματος είναι το σύνολο των υποψηφίων σημείων και οι ακμές του γραφήματος το σύνολο των συντομότερων μονοπατιών μεταξύ οποιωνδήποτε δύο γειτονικών υποψηφίων σημείων. Οι κόμβοι και οι ακμές είναι τα αποτελέσματα με βάση την χωρική / χρονική ανάλυση.

Result Matching: Αυτό το συστατικό αξιολογεί το υποψήφιο γράφημα χρησιμοποιώντας τα βάρη που αποδόθηκαν κατά την χωρική / χρονική ανάλυση. Ταιριάζει την δεδομένη τροχιά με το μονοπάτι με το υψηλότερο σκορ στο υποψήφιο γράφημα. Τα αποτελέσματα μπορεί να οπτικοποιηθούν στη συνέχεια, σε ένα περιβάλλον που μπορούν να προσαρμοστούν σε διαφορετικές συσκευές τελικών χρηστών. Τα αποτελέσματα μπορούν να αποθηκεύονται σε μια βάση δεδομένων κίνησης για την υποστήριξη εξωτερικών εφαρμογών, όπως η διαχείριση της κυκλοφορίας. [LZZX2009].

ΚΕΦΑΛΑΙΟ 3

Εύρεση Πιθανής Περιοχής Δραστηριοποίησης

3.1 Εισαγωγή

Το κεφαλαίο αυτό αναφέρεται στην εύρεση της πιθανής περιοχής δραστηριοποίησης (Potential Area of Activity PAA) εν κινήσει χρηστών. Για την εύρεση της πιθανής περιοχής δραστηριοποίησης εν κινήσει χρηστών μελετήθηκαν αναλυτικά οι εργασίες [PJ1999], [NVS2010]. Αναλυτική περιγραφή για τις συγκεκριμένες εργασίες έχει γίνει στο κεφάλαιο 2 όπου αναφέρεται όλο το θεωρητικό υπόβαθρο για την επίλυση του συγκεκριμένου προβλήματος. Σύμφωνα με την θεωρητική μελέτη που έχει παρουσιαστεί στις εργασίες αυτές υλοποιήθηκε η λύση του συγκεκριμένου προβλήματος. Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα με πολλές δυνατότητες, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση PostGIS. Το PostGIS είναι ένα ανοικτό πρόγραμμα λογισμικού που προσθέτει υποστήριξη για γεωγραφικά αντικείμενα στην αντικείμενο-σχεσιακή βάση δεδομένων PostgreSQL. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα QGIS.

3.2 Εργαλεία

Στην ενότητα αυτή παρουσιάζονται τα εργαλεία που χρησιμοποιήθηκαν για την υλοποίηση της παρούσας εργασίας. Το Σύστημα Διαχείρισης Βάσης Δεδομένων που χρησιμοποιήθηκε είναι η PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση του PostGIS. Το QGIS χρησιμοποιήθηκε για την οπτικοποίηση των αποτελεσμάτων. Στην συνέχεια, γίνεται μια συνοπτική αναφορά στα χαρακτηριστικά της PostgreSQL, του PostGIS και του QGIS.

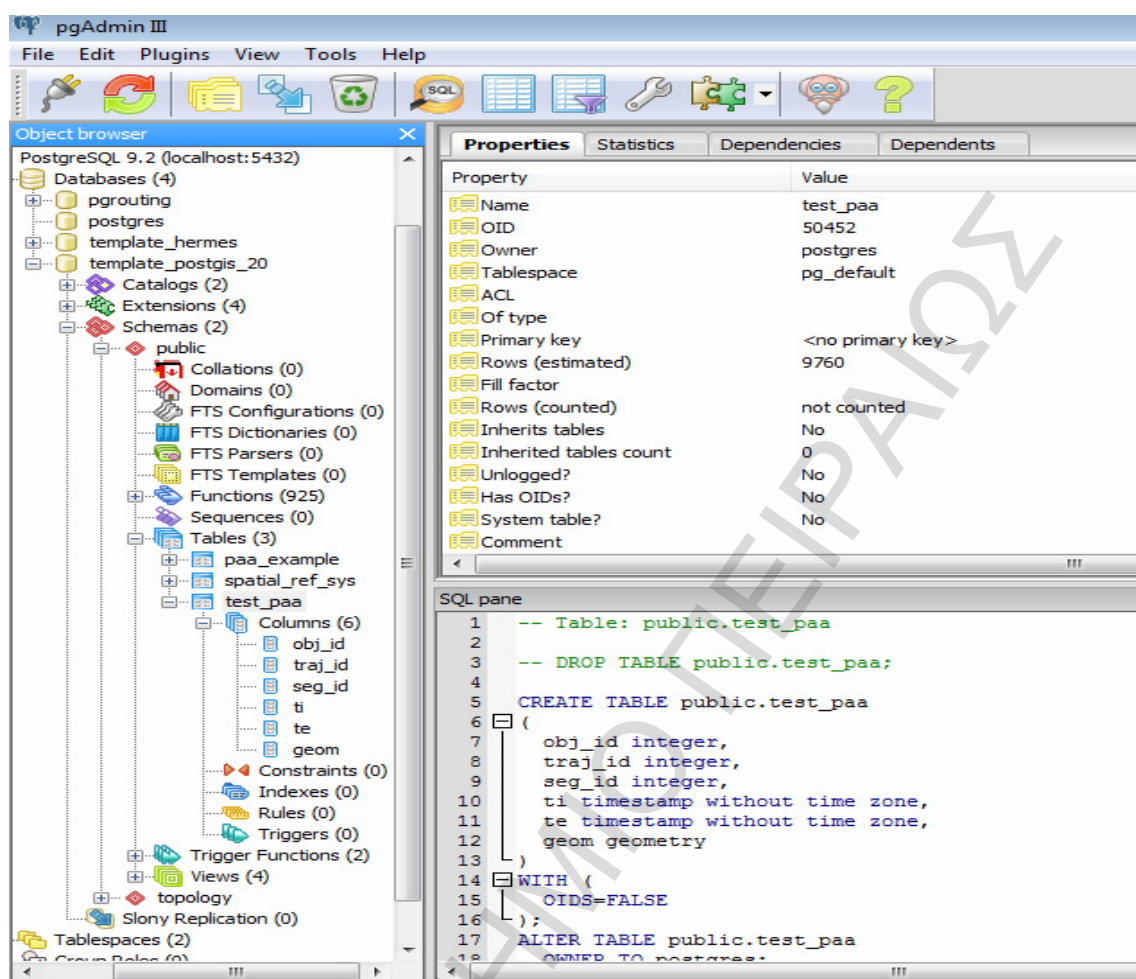
3.2.1 Η PostgreSQL

Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα, PostgreSQL. Η ανάπτυξη της διαρκεί πάνω από δύο δεκαετίες και βασίζεται σε μια αποδεδειγμένα καλή αρχιτεκτονική η οποία έχει δημιουργήσει μια ισχυρή αντίληψη των χρηστών της γύρω από την αξιοπιστία, την ακεραιότητα και την ορθή λειτουργία. Η PostgreSQL είναι ένα αντικείμενο-σχεσιακό σύστημα διαχείρισης βάσεων δεδομένων. Υποστηρίζει ένα μεγάλο μέρος του προτύπου SQL και προσφέρει πολλά σύγχρονα χαρακτηριστικά όπως: σύνθετα ερωτήματα, ξένα κλειδιά, triggers, views, transactional integrity. Επίσης, η PostgreSQL μπορεί να επεκταθεί από τον χρήστη, με την προσθήκη νέων τύπων δεδομένων, λειτουργιών, τελεστών, μεθόδων δεικτοδότησης, διαδικαστικών γλωσσών. Η PostgreSQL τρέχει σε όλα τα λειτουργικά συστήματα, το Linux, Unix, MacOS X και τα Windows. Είναι συμβατή με το ACID άρα οι συναλλαγές με την βάση δεδομένων είναι αξιόπιστες. Συμπεριλαμβάνει τους περισσότερους SQL92 και SQL99 τύπους δεδομένων. Βασίζεται στο μοντέλο πελάτη/εξυπηρετητή. Μπορεί να διαχειριστεί μεγάλο αριθμό ταυτόχρονων χρηστών καθώς και μεγάλο όγκο δεδομένων. Διαθέτει περιβάλλοντα προγραμματισμού για: C, C++, Java, Perl, Python, Ruby, Tcl. Έχει σχεδιαστεί για να ικανοποιήσει τις ανάγκες των χρηστών, από απλό γράψιμο sql έως την ανάπτυξη πολύπλοκων βάσεων δεδομένων. Είναι διαθέσιμο σε περισσότερες από δώδεκα γλώσσες.

Πλήρης υποστήριξη συναρτήσεων συγκεντρωτικών αποτελεσμάτων (GROUP_BY) όπως COUNT, SUM, AVG, MIN, MAX. Δυνατότητα δημιουργίας νέων συγκεντρωτικών συναρτήσεων. Υποστήριξη όλων των τύπων ενώσεων (cross, inner, outer, left, right, full, natural). Βιβλιοθήκη συναρτήσεων και τελεστών με ορισμένες προεγκατεστημένες συναρτήσεις όπως math, date/time, string, geometric, formatting κ.α.

Η PostgreSQL διαθέτει διάφορους τύπους δεδομένων: λογικοί, δυαδικοί, χαρακτήρες, αριθμητικοί, ημερομηνία/ώρα, γεωμετρικοί, δικτυακοί και συστήματος, επίσης και την δυνατότητα δημιουργίας νέων τύπων δεδομένων χρησιμοποιώντας την εντολή CREATE TYPE. Συναρτήσεις ορισμένες από τον χρήστη, οι οποίες μπορούν να γραφούν σε γλώσσες προγραμματισμού όπως C, SQL, PL/pgSQL, TCL, Perl, Python και Ruby. Προσωρινοί πίνακες. Το μέγεθος του πίνακα και της βάσης δεδομένων είναι σχεδόν απεριόριστο.

Η Διαχείριση της βάσης δεδομένων γίνεται μέσω του εργαλείου pgAdmin αλλά και με την χρήση εφαρμογών τρίτων όπως PgAccess, PhpPgAdmin, WinSQL. Η Σύνδεση με την βάση δεδομένων επιτυγχάνεται επιλέγοντας το εικονίδιο "add a connection to a server". Στην συνέχεια θα πρέπει να εισάγουμε όνομα χρήστη το οποίο είναι "postgres" καθώς και τον κωδικό πρόσβασης που επιλέξαμε κατά την διαδικασία εγκατάστασης.



Εικόνα 3-1: pgAdmin, το περιβάλλον διεπαφής της PostgreSQL

3.2.2 Το PostGIS

Το PostGIS είναι μια επέκταση της PostgreSQL ειδικά για χωρικά δεδομένα. Το PostGIS είναι ανοικτού κώδικα και η τελευταία έκδοση περιλαμβάνει υποστήριξη για διανυσματικά και δικτυωτά δεδομένα, χωρικά ευρετήρια, μετατροπή δεδομένων μεταξύ συστημάτων γεωγραφικών συντεταγμένων και λειτουργίες για ανάλυση και επεξεργασία των αντικειμένων GIS. Το PostGIS είναι ένα λογισμικό ανοικτού κώδικα ειδικά σχεδιασμένο για την διαχείριση χωρικών δεδομένων, χωρίς αδειοδότηση ή περιορισμούς χρήσης και υποστηρίζει γεωγραφικά αντικείμενα στην PostgreSQL. Επίσης υποστηρίζει την αναπαράσταση χωρικών οντοτήτων όχι μόνο 2-διαστατων αλλά και 3-διαστατων και 4-διαστατων γεωμετριών. Η εγκατάσταση του PostGIS γίνεται αυτόματα από το εργαλείο <<Application Stack Builder>> της PostgreSQL με την προϋπόθεση ότι πρώτα έχουμε εγκαταστήσει την βάση δεδομένων PostgreSQL.

Το PostGIS ακολουθεί τις λειτουργίες και προδιαγραφές του προτύπου Open Geospatial Consortium (OGC). Παρέχει ένα σύνολο από γεωμετρικούς τύπους δεδομένων, χωρικούς τελεστές, χωρικές συναρτήσεις σύμφωνα με το πρότυπο OGC και συστήματα αναφοράς (Spatial Reference System (SRID)). Ο χρήστης έχει την δυνατότητα ορισμού του συστήματος αναφοράς στα γεωμετρικά δεδομένα. Υπάρχουν πολλά συστήματα αναφοράς και όλα χαρακτηρίζονται από ένα μοναδικό κωδικό EPSG. Για παράδειγμα το σύστημα αναφοράς GRS87 με EPSG κωδικό 2100 και μονάδα μέτρησης τα μέτρα, το σύστημα αναφοράς WGS84 με EPSG κωδικό 4326 και μονάδα μέτρησης τις μίλλες. Είναι αποθηκευμένα στο πίνακα SPATIAL_REF_SYS του PostGIS.

Το PostGIS έχει ένα δυναμικό γεωμετρικό τύπο δεδομένων, τον *geometry*, τα αντικείμενα που περιέχουν γεωμετρικά δεδομένα ορίζονται με αυτό τον τύπο. Τα γεωγραφικά αντικείμενα που μπορούν να αποθηκευτούν στον τύπο *geometry* είναι τα *points* (σημεία), *linestrings* (γραμμές), *polygons* (πολύγωνα), *multipoints* (πολυσημεία), *multilinestrings* (πολύγραμμες), *multipartypolgons* (πολυπολύγωνο). Υπάρχουν δυο μέθοδοι αναπαράστασης χωρικών δεδομένων σύμφωνα με το OGS: ο *Well-known text (WKT)* και *Well-known binary (WKB)*. Η αναπαράσταση υπό την μορφή WKT είναι πιο κατανοητή στον άνθρωπο. Ακολουθούν κάποια παραδείγματα:

Σημείο: 'POINT(X Y)'

Γραμμή: 'LINESTRING(X1 Y1, X2 Y2, ...)'

Πολύγωνο: 'POLYGON((X1 Y1, X2 Y2, ..., X1 Y1), (Xn,Yn, Xn+1 Yn+1,..., Xn Yn))'

Το PostGIS υποστηρίζει χωρικές συναρτήσεις για τον προσδιορισμό χωρικών μετρήσεων, όπως η περιοχή (*area*), η απόσταση (*distance*), το μήκος (*length*) και η περίμετρος (*perimeter*). Υποστηρίζει χωρικές συναρτήσεις για τον προσδιορισμό χωρικών πράξεων, όπως ένωση (*union*), διαφορά (*difference*), συμμετρική διαφορά (*symmetric difference*) και *buffers*. Τέλος, το PostGIS παρέχει την δυνατότητα οπτικοποίησης δεδομένων μέσω ειδικών εφαρμογών, όπως το QuantumGIS, uDig, MapServer και GeoServer.

3.2.3 To Quantum GIS

Το QGIS είναι ανοιχτού κώδικα Σύστημα Γεωγραφικών Πληροφοριών που παρέχει προβολή, επεξεργασία και ανάλυση δεδομένων. Παρόμοια με άλλα GIS συστήματα, το λογισμικό QGIS επιτρέπει στους χρήστες να δημιουργήσουν χάρτες που χρησιμοποιούν διαφορετικές χαρτογραφικές προβολές. Οι χάρτες μπορούν να συναρμολογηθούν σε διαφορετικές μορφές και για διαφορετικές χρήσεις. Το QGIS επιτρέπει η απεικόνιση των χαρτών να γίνεται με δυο τρόπους: ως *raster* ή *vector* επιπέδων. Τυπικά για αυτό το είδος του λογισμικού τα διανυσματικά δεδομένα αποθηκεύονται ως σημεία, γραμμές ή πολύγωνα. Οι *raster* εικόνες υποστηρίζονται και το λογισμικό μπορεί να εκτελέσει γεωαναφορά των εικόνων. Ο χρήστης έχει την δυνατότητα να δημιουργεί, να επεξεργαστεί, να απεικονίσει, να αναλύσει και να δημοσιεύσει γεωχωρικές πληροφορίες. Το QGIS παρέχει ενοποίηση με άλλα GIS πακέτα ανοιχτού κώδικα, συμπεριλαμβανομένης του PostGIS, GRASS και MapServer για να δώσει στους χρήστες εκτεταμένη λειτουργικότητα. Υπάρχουν *plugins* για την υπηρεσία Google Geocoding, για την διασύνδεση με την PostgreSQL και MySQL βάσεις δεδομένων. Το QGIS τρέχει σε πολλά λειτουργικά συστήματα συμπεριλαμβανομένου του MacOS X, Linux, UNIX και Windows.

3.3 Potential Area of Activity (PAA)

Τα ζητούμενα είναι:

Περίπτωση 1:

Ποια είναι η πιθανή περιοχή δραστηριότητας PAA του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i = t_x$ και $t_y = t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για όλη την τροχιά του.

Περίπτωση 2:

Ποια είναι η πιθανή περιοχή δραστηριότητας PAA του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i < t_x < t_y < t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για ένα ενδιάμεσο χρονικό διάστημα, δηλαδή, από t_x έως t_y .

Περίπτωση 3:

Ποια είναι η πιθανή περιοχή δραστηριότητας PAA του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i < t_x = t_y < t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για ένα χρονικό διάστημα της κίνησής του, αφού $t_x = t_y$.

Στη συνέχεια θα γίνει η αναλυτική περιγραφή της συνάρτησης PAA, η οποία βρίσκει τη πιθανή περιοχή δραστηριότητας ενός κινούμενου αντικειμένου. Η συνάρτηση PAA περιλαμβάνει

και τις τρεις περιπτώσεις που περιγράφηκαν παραπάνω. Η συνάρτηση αυτή δέχεται σαν είσοδο την μέγιστη ταχύτητα που μπορεί να κινείται το αντικείμενο, το id της τροχιάς, την χρονική στιγμή t_x στο χρονικό διάστημα κίνησής του, ισχύει $t_i \leq t_x \leq t_e$ και την χρονική στιγμή t_y στο χρονικό διάστημα κίνησης του, ισχύει $t_i \leq t_y \leq t_e$.

Παρακάτω δίνεται ο ψευδοκώδικας της συνάρτησης (PAA) εύρεσης πιθανής περιοχής δραστηριότητας. Όπως, φαίνεται στην Εικόνα 3-2, στην πρώτη γραμμή ορίζεται η συνάρτηση PAA και τα ορίσματά της όπως, περιγράψαμε και παραπάνω είναι η μέγιστη ταχύτητα του αντικειμένου, τα id της τροχιάς, δυο χρονικά διαστήματα t_x και t_y στο χρονικό διάστημα κίνησής του. Στη συνέχεια υπολογίζονται τα στοιχεία της έλλειψης.

Άρα, υπολογίζονται τα στοιχεία της έλλειψης ως εξής: από τη γραμμή 4 έως την γραμμή 9 υπολογίζεται η απόσταση ανάμεσα στα δυο χρονικά διαστήματα t_x και t_y , η χρονική απόσταση ανάμεσα στα δυο χρονικά διαστήματα t_x και t_y , το x του πρώτου σημείου, το y του πρώτου σημείου και η κλίση. Στη συνέχεια από την γραμμή 10 έως την γραμμή 29 μέσα στην εντολή επανάληψης υπολογίζονται οι άξονες της κάθε έλλειψης και τα τέσσερα σημεία της κάθε έλλειψης. Πρώτα υπολογίζεται ο μεγάλος άξονας και μετά ο μικρός άξονας της κάθε έλλειψης. Από την γραμμή 16 έως την γραμμή 24 υπολογίζονται τα 4 σημεία της κάθε έλλειψης. Τα τέσσερα σημεία της έλλειψης είναι τα m_1 , m_2 , m_3 και m_4 . Τα σημεία αυτά φαίνονται στην Εικόνα 3-3. Στις γραμμές 16 έως 24 φαίνεται ο τρόπος υπολογισμού των σημείων αυτών. Στην συνέχεια, αφού έχουμε υπολογίσει τα τέσσερα σημεία της έλλειψης τα εισάγουμε στο βοηθητικό πίνακα `foo[]` για να υπολογίσουμε το παραλληλόγραμμο που περικλείει την έλλειψη.

```

1  Function PAA (v_max, trajectory_id, tx, ty)
2  Begin
3  i := 1;
4  // Υπολογισμός Έλλειψης
5  ds[] = η απόσταση ανάμεσα στα tx και ty ;
6  dt[] = η χρονική απόσταση ανάμεσα στα tx και ty ;
7  x1[] = το x του πρώτου σημείου ανάμεσα στα tx και ty;
8  y1[] = το y του πρώτου σημείου ανάμεσα στα tx και ty;
9  theta[] = η κλίση ανάμεσα στα tx και ty;
10 FOR r in
11 LOOP
12     // Υπολογισμός μεγάλου άξονα της έλλειψης
13     major_axis = v_max * dt[i];
14     // Υπολογισμός μικρού άξονα της έλλειψης
15     minor_axis = sqrt (major_axis ^ 2 - ds[i] ^ 2) ;
16     // Υπολογισμός των 4 σημείων της έλλειψης
17     // Ορίζω το σημείο M1 : (X1+c-a , Y1-b)
18     m1 = ((x1[i] + (ds[i]/2) - (major_axis/2), y1[i] - (minor_axis/2)) ;
19     // Ορίζω το σημείο M2 : (X1+c-a , Y1+b)
20     m2 = ((x1[i] + (ds[i]/2) - (major_axis/2), y1[i] + (minor_axis/2)) ;
21     // Ορίζω το σημείο M3 : (X1+c+a , Y1+b)
22     m3 = ((x1[i] + (ds[i]/2) + (major_axis/2), y1[i] + (minor_axis/2)) ;
23     // Ορίζω το σημείο M4 : (X1+c+a , Y1-b)
24     m4 = ((x1[i] + (ds[i]/2) + (major_axis/2), y1[i] - (minor_axis/2)) ;

```

```

25      --Εισαγωγή 4 σημείων της κάθε έλλειψης σε βοηθητικό πίνακα για να
26      -- υπολογίσουμε το παραλληλόγραμμα που περικλείει την έλλειψη
27      foo [i] = m1, m2, m3, m4 ;
28      i := i + 1;
29  END LOOP;
30  --Υπολογίζουμε τα παραλληλόγραμμα των ελλείψεων και τα εισάγουμε στο array ellipse
31  ellipse[i]=υπολογίζουμε τα παραλληλόγραμμα της κάθε έλλειψης από τα σημεία του πίνακα foo[i] ;
32  // Υπολογισμός κύκλων Ci, Ce
33  // Υπολογισμός της ακτίνας r1 του κύκλου Ci
34  r1 = v_max * (ty - ti) ;
35  // Υπολογισμός της ακτίνας r2 του κύκλου Ce
36  r2 = v_max * (te - tx) ;
37  Ci = δημιουργία του κύκλου Ci ;
38  Ce = δημιουργία του κύκλου Ce ;
39  // Υπολογισμός PAA
40  --Υπολογισμός της τομής του Ci με την πρώτη (χρονικά) έλλειψη
41  C1= τομή (Ci, ellipse[1]);
42  --Υπολογισμός της τομής του Ce με την τελευταία (χρονικά) έλλειψη
43  C2= τομή (Ce, ellipse[i-1]);
44  UnionOfEllipses = ellipse [2];
45  FOR j in 3..i-2
46  LOOP
47      --Υπολογισμός της ένωσης των ελλείψεων από έλλειψη 2 έως i-1
48      UnionOfEllipses = ένωση (UnionOfEllipses, ellipse[j]);
49  END LOOP;
50  --Υπολογισμός της ένωσης του UnionOfEllipses και των κύκλων C1 και C2
51  PAA = ένωση (ένωση (C1, UnionOfEllipses), C2);
52  return PAA;
53  end;

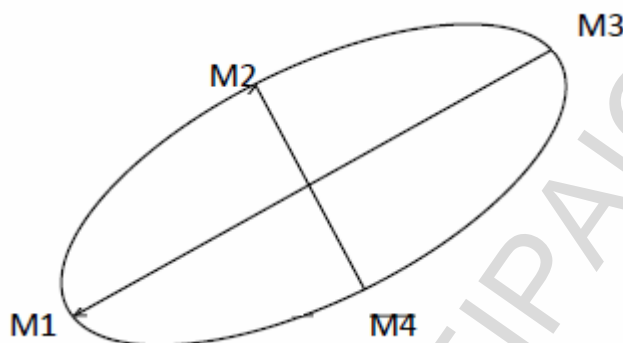
```

Εικόνα 3-2: Ψευδοκώδικας της συνάρτησης PAA

Στη συνέχεια θα πρέπει να γίνει ο υπολογισμός των ελλείψεων. Στις γραμμές 30, 31 υπολογίζουμε τα παραλληλόγραμμα των ελλείψεων και τα εισάγουμε στο πίνακα ellipse[]. Υπολογίζουμε το παραλληλόγραμμα που περικλείει την κάθε έλλειψη από τα σημεία του πίνακα foo[i] και μετά τα αποτελέσματα τα εισάγουμε στο πίνακα ellipse[]. Στη συνέχεια υπολογίζουμε τους δυο κύκλους. Από την γραμμή 32 έως την γραμμή 38 υπολογίζονται οι κύκλοι αυτοί. Πρώτα υπολογίζουμε την ακτίνα r1 του κύκλου Ci, μετά την ακτίνα r2 του κύκλου Ce, όπως φαίνεται στις γραμμές 34 και 36 αντίστοιχα. Μετά υπολογίζεται ο κύκλος Ci και στη συνέχεια ο κύκλος Ce, όπως φαίνεται στις γραμμές 37 και 38 αντίστοιχα.

Τέλος, από την γραμμή 39 έως την γραμμή 53 υπολογίζεται το PAA δηλαδή, η πιθανή περιοχή δραστηριότητας του κινούμενου αντικειμένου. Πρώτα υπολογίζουμε το κύκλο C1 από την τομή του Ci με την πρώτη (χρονικά) έλλειψη, στην γραμμή 41. Μετά υπολογίζουμε το κύκλο C2 από την τομή του Ce με την τελευταία (χρονικά) έλλειψη, στην γραμμή 43. Στην γραμμή 44

καταχωρείται στην μεταβλητή UnionOfEllipses η δεύτερη έλλειψη. Μετά στην επανάληψη από την γραμμή 45 έως την γραμμή 49 υπολογίζεται η ένωση των ελλείψεων από την έλλειψη 2 έως την έλλειψη i-1. Τέλος, η πιθανή περιοχή δραστηριότητας PAA του κινούμενου αντικείμενου είναι η ένωση των UnionOfEllipses και των κύκλων C1 και C2. Στην γραμμή 51 υπολογίζεται το PAA και το αποτέλεσμα καταχωρείται στην μεταβλητή PAA. Επιστρέφεται το αποτέλεσμα του PAA και εδώ τελειώνει η συνάρτηση εύρεσης πιθανής περιοχής δραστηριότητας κινούμενου αντικείμενου.



Εικόνα 3-3: Τα σημεία της έλλειψης

3.4 Περιγραφή Υλοποίησης της PAA

Τα χωροχρονικά δεδομένα βρίσκονται σε ένα csv αρχείο, με τιμές για τα κάτωθι μεγέθη :

| Πεδίο | Περιγραφή | Τύπος |
|---------|--|-----------------------------|
| obj_id | Το id του κινούμενου αντικείμενου | integer |
| traj_id | Το id της τροχιάς | integer |
| seg_id | Το id του segment | integer |
| ti | Η χρονική στιγμή έναρξης για μια συγκεκριμένη τροχιά | timestamp without time zone |
| te | Η χρονική στιγμή τέλους στη συγκεκριμένη τροχιά | timestamp without time zone |
| geom | Μια linestring γεωμετρία | geometry |

Πίνακας 3-1: Περιγραφή Δεδομένων

Έγιναν οι πιο κάτω διαδικασίες :

Δημιουργήθηκε ένας πίνακας για τα χωροχρονικά δεδομένα που δίνονται, στη Σχεσιακή Βάση Δεδομένων Postgresql η οποία εμπεριέχει και το σύστημα διαχείρισης χωροχρονικών δεδομένων, το PostGIS.

Δόθηκε η SQL εντολή:

```
CREATE TABLE test_paa
(
  obj_id integer,
  traj_id integer,
  seg_id integer,
```

```

ti timestamp without time zone,
te timestamp without time zone,
geom geometry
);

```

Και στη συνέχεια για το φόρτωμα των δεδομένων από το csv αρχείο C:\test_data, στον πίνακα test_paa που δημιουργήθηκε παραπάνω, δόθηκε η παρακάτω SQL εντολή:

```

COPY test_paa (obj_id, traj_id, seg_id, ti, te, geom)
FROM 'C:\test_data'
WITH DELIMITER ',';
CSV HEADER

```

Στη συνέχεια κατασκευάστηκε η stored function PAA η οποία υλοποιεί τα 3 ζητούμενα.

| Είσοδος | | |
|---------------|--|-----------------------------|
| Παράμετρος | Όνομα Παραμέτρου | Τύπος Παραμέτρου |
| v_max | Η μέγιστη ταχύτητα | float |
| trajectory_id | Το id της τροχιάς | int |
| tx | Χρονική στιγμή στο χρονικό διάστημα κίνησής του Ισχύει $t_i \leq tx \leq t_e$ | timestamp without time zone |
| ty | Χρονική στιγμή στο χρονικό διάστημα κίνησής του Ισχύει $t_i \leq ty \leq t_e$ | timestamp without time zone |

Πίνακας 3-2: Παράμετροι της συνάρτησης PAA

| Εξοδος |
|---|
| Η function επιστρέφει δεδομένα τύπου geometry |

| Βοηθητικοί πίνακες | |
|--------------------|--------------------------------|
| foo | Γίνεται drop μετά τη χρήση του |

| Συναρτήσεις του PostGIS που χρησιμοποιήθηκαν |
|--|
| ST_Distance |
| ST_PointN |
| Extract epoch from |
| ST_X |

| |
|-----------------|
| ST_Y |
| ST_MakePoint |
| ST_Azimuth |
| ST_Envelope |
| ST_Collect |
| ST_Buffer |
| ST_SetSRID |
| ST_Intersection |
| ST_Union |

Πίνακας 3-3: Συναρτήσεις του PostGIS που χρησιμοποιήθηκαν

3.5 Οπτικοποίηση Αποτελεσμάτων

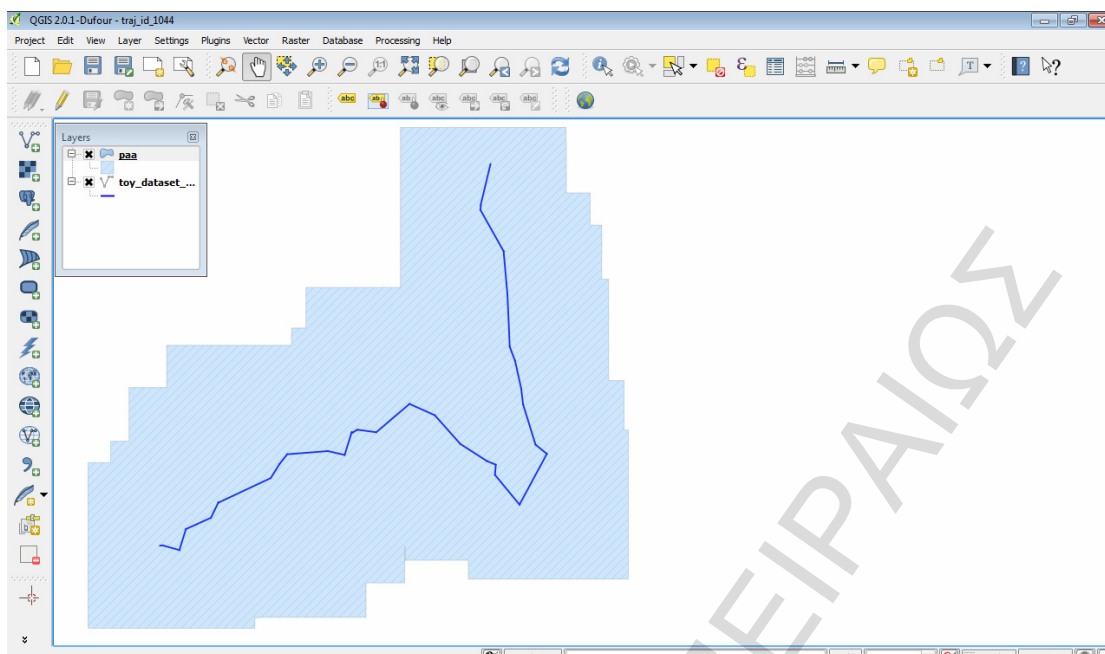
Τα αποτελέσματα για την εύρεση της πιθανής περιοχής δραστηριοποίησης έχουν οπτικοποιηθεί στο QGIS. Η διαδικασία που γίνεται είναι η εξής: Το QGIS δίνει την δυνατότητα σύνδεσης με την PostgreSQL. Μέσω της επιλογής “Add PostGIS Layers” δημιουργούμε στην αρχή μια καινούργια σύνδεση δηλαδή μια “Create a new PostGIS connection” δίνοντας όλες τις πληροφορίες της σύνδεσης (Name, Service, Host, Port, Database, SSL Mode, Username, Password). Αφού δημιουργήθηκε η PostGIS σύνδεση, μετά επιλέγουμε το name που δόθηκε πριν και πατώντας το κουμπί «connect» γίνεται η σύνδεση και μπορούμε να προσθέσουμε το πίνακα της βάσης δεδομένων που μας ενδιαφέρει. Αφού επιλέξουμε το πίνακα και πατώντας το κουμπί «add» βλέπουμε στο παράθυρο με τα layers να έχει γίνει η πρόσθεση ενός καινούργιου επιπέδου με το όνομα του πίνακα που επιλέξαμε.

Στην συνέχεια παρουσιάζονται δυο παραδείγματα από δυο διαφορετικές τροχιές κίνησης του αντικειμένου και για τις τρεις περιπτώσεις που περιγράψαμε παραπάνω. Τα αποτελέσματα της συνάρτησης PAA αποθηκεύονται στον πίνακα paa στην βάση δεδομένων. Αυτό που παρατηρείται είναι ότι για το ίδιο trajectory και χρονικά διαστήματα αν δώσουμε διαφορετική μέγιστη ταχύτητα έχουμε διαφορετική περιοχή δραστηριοποίησης δηλαδή, για το ίδιο trajectory και χρονικά διαστήματα αν η μέγιστη ταχύτητα είναι μικρή τότε η πιθανή περιοχή δραστηριοποίησης είναι πιο μικρή σε σχέση με μια μεγαλύτερη μέγιστη ταχύτητα.

3.5.1 Παράδειγμα 1

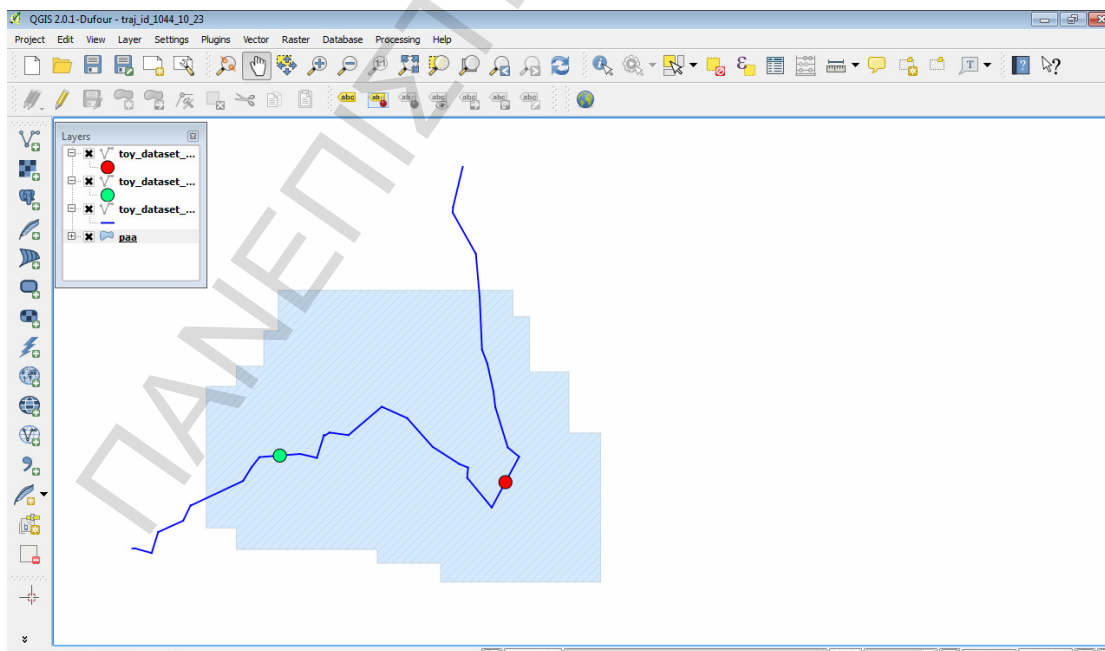
Στο πρώτο παράδειγμα θα βρούμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για μια τροχιά κίνησής του. Όπως, περιγράψαμε παραπάνω έχουμε τρεις περιπτώσεις: Στην περίπτωση 1, οπτικοποιείται η τροχιά του κινούμενου αντικειμένου και η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για όλη την τροχιά του. Στην περίπτωση 2, οπτικοποιείται η τροχιά του κινούμενου αντικειμένου, η χρονική στιγμή t_x , η χρονική στιγμή t_y και η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για t_x έως t_y . Στην περίπτωση 3, οπτικοποιείται η τροχιά του κινούμενου αντικειμένου, η χρονική στιγμή $t_x = t_y$ και η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για $t_x = t_y$.

Περίπτωση 1: Η πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i = t_x$ και $t_j = t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για όλη την τροχιά του. Στην παρακάτω εικόνα η μπλε γραμμή είναι η τροχιά κίνησης του αντικειμένου. Η γαλάζια περιοχή είναι η πιθανή περιοχή δραστηριοποίησης του αντικειμένου.



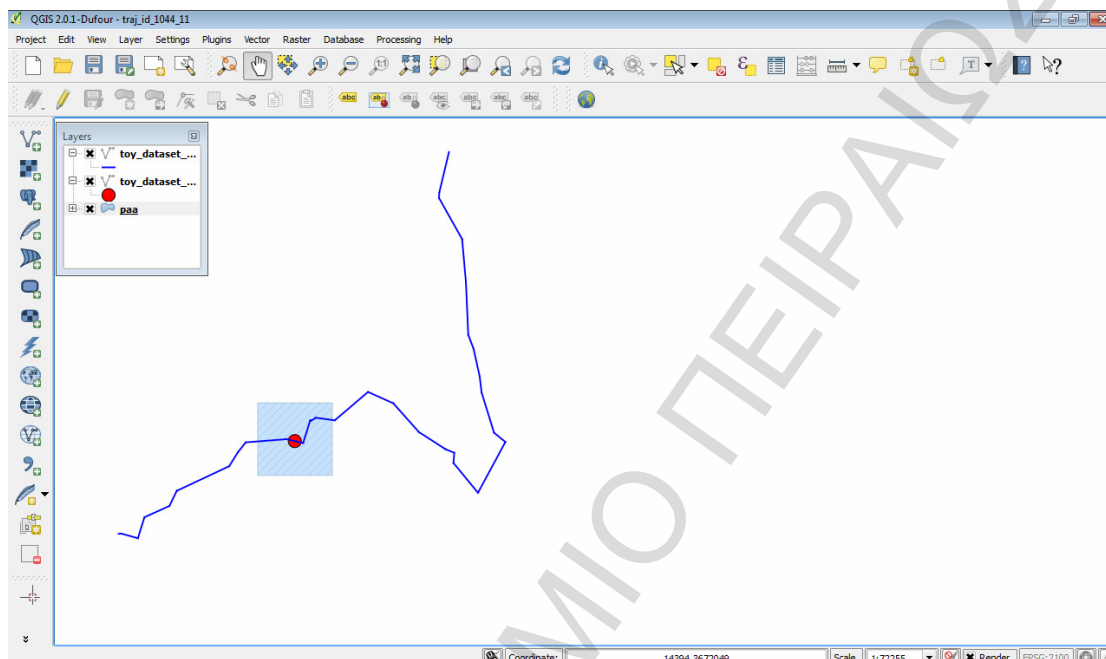
Εικόνα 3-4: Παράδειγμα_1_ Περίπτωση_1

Περίπτωση 2: Η πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικείμενου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητα που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i < t_x < t_y < t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικείμενου για ένα ενδιάμεσο χρονικό διάστημα, δηλαδή, από t_x έως t_y . Στην παρακάτω εικόνα η μπλε γραμμή είναι η τροχιά κίνησης του αντικείμενου. Η πράσινη κουκίδα είναι η χρονική στιγμή t_x και η κόκκινη κουκίδα η χρονική στιγμή t_y . Η γαλάζια περιοχή είναι η πιθανή περιοχή δραστηριοποίησης του αντικείμενου σύμφωνα με το ενδιάμεσο χρονικό διάστημα, t_x έως t_y .



Εικόνα 3-5: Παράδειγμα_1_ Περίπτωση_2

Περίπτωση 3: Η πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i < t_x = t_y < t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για ένα χρονικό διάστημα της κίνησής του, αφού $t_x = t_y$. Στην παρακάτω εικόνα η μπλε γραμμή είναι η τροχιά κίνησης του αντικειμένου. Η κόκκινη κουκίδα είναι η χρονική στιγμή t_x και η χρονική στιγμή t_y , αφού $t_x = t_y$. Η γαλάζια περιοχή είναι η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για ένα χρονικό διάστημα, αφού $t_x = t_y$.

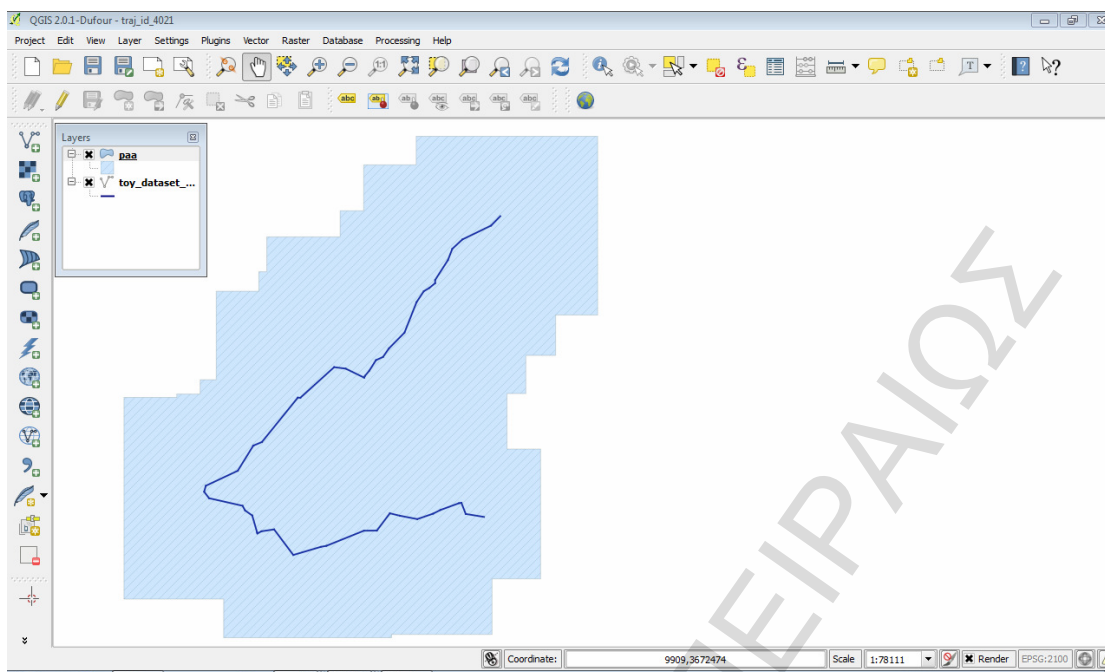


Εικόνα 3-6: Παράδειγμα_1_ Περίπτωση_3

3.5.2 Παράδειγμα 2

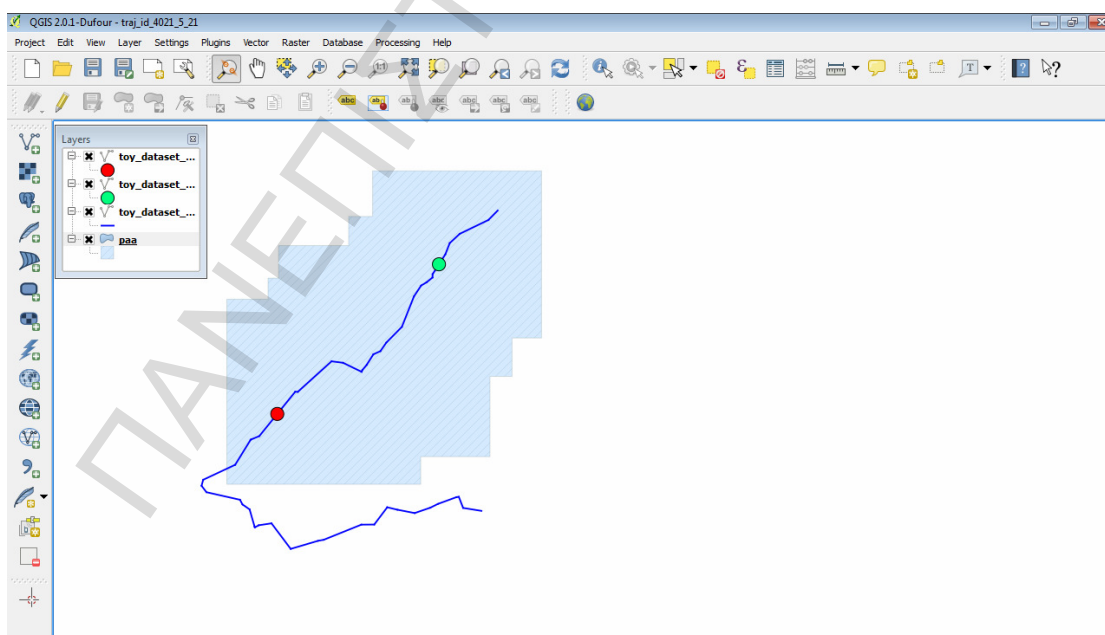
Στο δεύτερο παράδειγμα θα βρούμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για άλλη τροχιά κίνησής του. Όπως, περιγράψαμε παραπάνω έχουμε τρεις περιπτώσεις: Στην περίπτωση 1, οπτικοποιείται η τροχιά του κινούμενου αντικειμένου και η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για όλη την τροχιά του. Στην περίπτωση 2, οπτικοποιείται η τροχιά του κινούμενου αντικειμένου, η χρονική στιγμή t_x , η χρονική στιγμή t_y και η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για t_x έως t_y . Στην περίπτωση 3, οπτικοποιείται η τροχιά του κινούμενου αντικειμένου, η χρονική στιγμή $t_x = t_y$ και η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για $t_x = t_y$.

Περίπτωση 1: Η πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i = t_x$ και $t_y = t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για όλη την τροχιά του. Στην παρακάτω εικόνα η μπλε γραμμή είναι η τροχιά κίνησης του αντικειμένου. Η γαλάζια περιοχή είναι η πιθανή περιοχή δραστηριοποίησης του αντικειμένου.



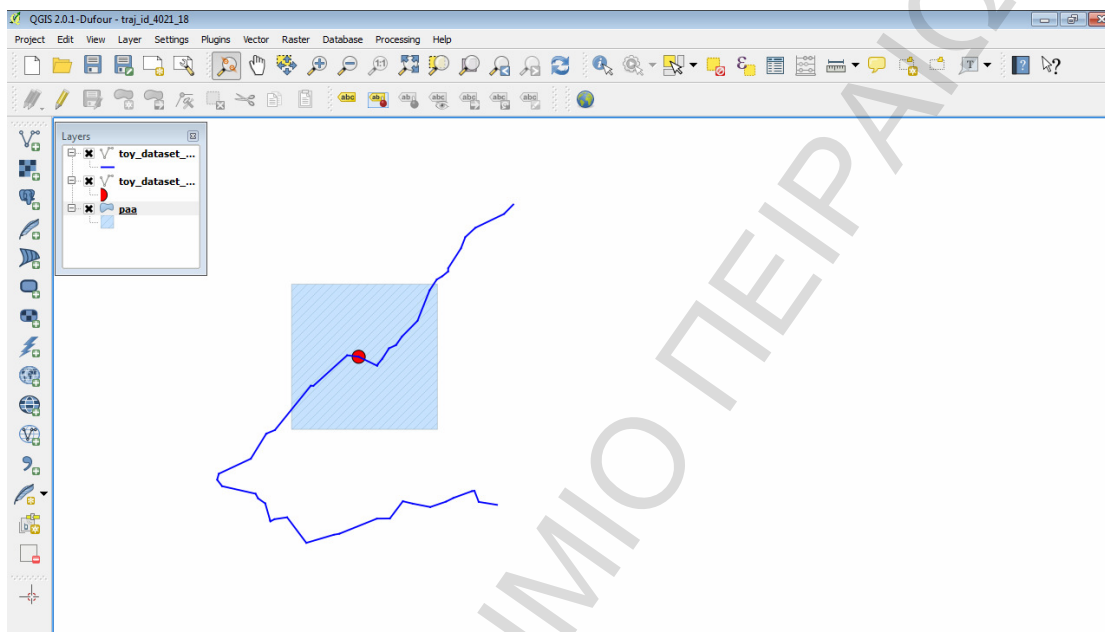
Εικόνα 3-7: Παράδειγμα_2_ Περίπτωση_1

Περίπτωση 2: Η πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i < t_x < t_y < t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για ένα ενδιάμεσο χρονικό διάστημα, δηλαδή, από t_x έως t_y . Στην παρακάτω εικόνα η μπλε γραμμή είναι η τροχιά κίνησης του αντικειμένου. Η πράσινη κουκίδα είναι η χρονική στιγμή t_x και η κόκκινη κουκίδα η χρονική στιγμή t_y . Η γαλάζια περιοχή είναι η πιθανή περιοχή δραστηριοποίησης του αντικειμένου σύμφωνα με το ενδιάμεσο χρονικό διάστημα, t_x έως t_y .



Εικόνα 3-8: Παράδειγμα_2_ Περίπτωση_2

Περίπτωση 3: Η πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου με δεδομένη την τροχιά κίνησης, την μέγιστη ταχύτητας που μπορεί να κινείται το αντικείμενο, δυο χρονικά διαστήματα t_x και t_y όπου, $t_i < t_x = t_y < t_e$. Στην περίπτωση αυτή υπολογίζουμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικειμένου για ένα χρονικό διάστημα της κίνησής του, αφού $t_x = t_y$. Στην παρακάτω εικόνα η μπλε γραμμή είναι η τροχιά κίνησης του αντικειμένου. Η κόκκινη κουκίδα είναι η χρονική στιγμή t_x και η χρονική στιγμή t_y , αφού $t_x = t_y$. Η γαλάζια περιοχή είναι η πιθανή περιοχή δραστηριοποίησης του αντικειμένου για ένα χρονικό διάστημα, αφού $t_x = t_y$.



Εικόνα 3-9: Παράδειγμα_2_ Περίπτωση_3

ΚΕΦΑΛΑΙΟ 4

To Map Matching

4.1 Εισαγωγή

Το κεφαλαίο αυτό αναφέρεται στο πρόβλημα του map-matching το οποίο είναι μια διαδικασία που έχει στόχο να προσεγγίζει όσο πιο πιστά γίνεται την τροχιά ενός κινούμενου αντικείμενου πάνω σε έναν χάρτη. Το πρόβλημα που ονομάζεται map-matching είναι η προσπάθεια να «ταιριάξουμε» τη εκτιμωμένη θέση με μια καμπύλη στο χάρτη και να εκτιμηθεί με μεγαλύτερη ακρίβεια η πραγματική θέση του χρήστη. Άρα, με την βοήθεια της διαδικασίας map-matching μεταφέρουμε τις εκτιμήσεις στο οδικό δίκτυο και ως αποτέλεσμα έχουμε μια καλύτερη εικόνα της τροχιάς που ακολουθήθηκε. Υπάρχει ένας αριθμός από διαφορετικούς τρόπους για να προσεγγίσει το map matching πρόβλημα. Μελέτη πάνω στο συγκεκριμένο πρόβλημα έχει γίνει στις εργασίες [BK1996], [WBK2000], [ZG2006], [LZZX2009], [AWB2011].

Για την υλοποίηση έχει χρησιμοποιηθεί η σχεσιακή βάση δεδομένων ανοικτού κώδικα με πολλές δυνατότητες, PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση του PostGIS. Το PostGIS είναι ένα ανοικτό πρόγραμμα λογισμικού που προσθέτει υποστήριξη για γεωγραφικά αντικείμενα στην αντικείμενο - σχεσιακή βάση δεδομένων PostgreSQL. Επίσης, έχει χρησιμοποιηθεί το pgRouting που είναι επέκταση του PostGIS και της PostgreSQL και δίνει δυνατότητες δρομολόγησης σε γεωγραφικά δεδομένα και άλλες λειτουργίες ανάλυσης του δικτύου. Το pgRouting δίνει την δυνατότητα χωρικών λειτουργιών σε δίκτυα. Για την εύρεση των αποτελεσμάτων έχουν χρησιμοποιηθεί οι συναρτήσεις του pgRouting. Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα QGIS.

4.2 Εργαλεία

Στην ενότητα αυτή παρουσιάζονται τα εργαλεία που χρησιμοποιήθηκαν για την υλοποίηση αυτής της ενότητας της παρούσας εργασίας. Το Σύστημα Διαχείρισης Χωρικής Βάσης Δεδομένων που χρησιμοποιήθηκε για τη μοντελοποίηση του δικτύου σε χωρική βάση δεδομένων είναι η PostgreSQL, εμπλουτισμένη με τη χωρική επέκταση PostGIS και τις συναρτήσεις του pgRouting οι οποίες υλοποιούν διαδικασίες δρομολόγησης σε δίκτυα. Ενώ το QGIS χρησιμοποιήθηκε για την οπτικοποίηση των αποτελεσμάτων. Στην συνέχεια, γίνεται μια συνοπτική αναφορά στα βασικά χαρακτηριστικά του pgRouting, του osm2po, καθώς συνοπτική αναφορά στα χαρακτηριστικά της PostgreSQL, του PostGIS και του QGIS, έχει γίνει ήδη στο κεφάλαιο 3 .

4.2.1 Το pgRouting

Το pgRouting είναι επέκταση του PostGIS και της PostgreSQL και δίνει δυνατότητες δρομολόγησης σε γεωγραφικά δεδομένα και άλλες λειτουργίες ανάλυσης του δικτύου. Επίσης το pgRouting δίνει την δυνατότητα χωρικών λειτουργιών σε δίκτυα. Το pgRouting έχει τη δυνατότητα οπτικοποίησης των χωρικών δεδομένων είτε μέσω των διεπαφών JDBC και ODBC είτε μέσω ερωτημάτων PL/pgSQL, σε διάφορα περιβάλλοντα απεικόνισης, όπως είναι οι ειδικές εφαρμογές QuantumGIS ή uDIG. Η δρομολόγηση γίνεται είτε σε υπολογιστές είτε σε φορητές συσκευές. Το κόστος μιας διαδρομής υπολογίζεται δυναμικά μέσω της SQL και η τιμή του κόστους μπορεί να προκύπτει συνδυάζοντας πολλά πεδία ενός πίνακα ή συνδυάζοντας πεδία από διαφορετικούς πίνακες της βάσης δεδομένων. Το pgRouting περιέχει συναρτήσεις στις οποίες έχουν υλοποιηθεί διάφοροι αλγόριθμοι δρομολόγησης για την εύρεση διαδρομής.

4.2.2 Το Osm2po

Το osm2po είναι μια εφαρμογή που εκτελεί δυο λειτουργίες. Μετατρέπει τα αρχεία osm σε αρχεία sql. Άρα, η εφαρμογή αυτή δέχεται ένα αρχείο osm και το μετατρέπει σε αρχείο sql, έτσι ώστε να έχουμε την δυνατότητα να εισάγουμε τα δεδομένα στην βάση δεδομένων με δυνατότητες δρομολόγησης. Το osm2po είναι και ένας μετατροπέας και μια μηχανή δρομολόγησης. Το osm2po είναι μετατροπέας που αναλύει XML-Data του OpenStreetMap και καθιστά δυνατό να δρομολογηθούν. Το osm2po είναι σε θέση να μετατρέψει μεγάλα σύνολα όπως europe.osm. Το osm2po δημιουργεί SQL αρχεία για το PostGIS, συμβατά με το pgRouting και Quantum GIS. Το osm2po τρέχει σε περιβάλλον Windows, Unix και Mac. Απαιτείται η Java 6 + να εγκατασταθεί. Ο πυρήνας του osm2po είναι μόνο ένα αρχείο. Δεν χρειάζεται εγκατάσταση. Η εκτέλεση γίνεται από την γραμμή εντολών.

4.3 Περιγραφή Map Matching Συναρτήσεων

Η παρακάτω συνάρτηση βρίσκει το πλησιέστερο λινκ σε δεδομένο κόμβο εντός της δεδομένης απόστασης. Έχει τρία ορίσματα: το `point` που είναι η αναπαράσταση του σημείου (κόμβου), το `distance` όπου η συνάρτηση θα αναζητήσει για ένα λινκ εντός αυτής της απόστασης, `table` όπου είναι το όνομα του πίνακα με τα δεδομένα. Η συνάρτηση αυτή αρχικά παίρνει το `x` και το `y` του σημείου που έχει δοθεί δηλαδή του `point`. Στη συνέχεια ψάχνει για ένα λινκ εντός της απόστασης `distance`. Αν βρέθηκε το λινκ επιστρέφει το `id` του διαφορετικά αν τα δεδομένα δεν μπορούν να συνδυαστούν επιστρέφει `NULL`.

```

1  Function Εύρεση_κοντινότερου_λινκ_εντός_απόστασης (point, distance, table)
2  Begin
3      -- Λήψη του x και y του σημείου
4      x:= row.x;
5      y:= row.y;
6      -- Ψάχνοντας για ένα λινκ εντός της απόστασης
7      For row in 'select id, distance (geom, point) from table where 'x-distance y-distance', 'x+distance y+distance' '
8      End Loop;
9      If row.id is NULL Then
10         -- 'Τα δεδομένα δεν μπορούν να συνδυαστούν';
11         Return NULL;
12     End If;
13     Return row.id;
14 End;
```

Η παρακάτω συνάρτηση βρίσκει το πλησιέστερο κόμβο σε δεδομένο κόμβο εντός της δεδομένης απόστασης. Έχει τρία ορίσματα: το `point` που είναι η αναπαράσταση του σημείου (κόμβου), το `distance` όπου η συνάρτηση θα αναζητήσει για ένα κόμβο εντός αυτής της απόστασης, `table` όπου είναι το όνομα του πίνακα με τα δεδομένα. Η συνάρτηση αυτή αρχικά παίρνει το `x` και το `y` του σημείου που έχει δοθεί δηλαδή του `point`. Στη συνέχεια ψάχνει για ένα κόμβο εντός της απόστασης `distance`. Αν βρέθηκε ο κόμβος επιστρέφει το `source` ή το `target` του διαφορετικά επιστρέφει `NULL`.

```

1  Function Εύρεση_κοντινότερου_κόμβου_εντός_απόστασης (point, distance, table)
2  Begin
3      -- Λήψη του x και y του σημείου
4      x:= row.x;
5      y:= row.y;
6      -- Λήψη κοντινότερου source
7      For row in 'select source, distance (StartPoint (geom), point) from table where
8          'x-distance y-distance', 'x+distance y+distance' '
9      End Loop;
10     d1:= row.dist;
11     source:= row.source;
12     -- Λήψη κοντινότερου target
13     For row in 'select target, distance (EndPoint (geom), point) from table where
14         'x-distance y-distance', 'x+distance y+distance' '
15     End Loop;
16     d2:= row.dist;
17     target:= row.target;
```

```

18 -- Έλεγχος τι είναι πιο κοντά - source ή target
19 IF d1<d2 THEN
20     node := source; d := d1;
21 ELSE
22     node := target; d := d2;
23 END IF;
24 IF d=NULL OR d>distance THEN
25     node := NULL;
26 END IF;
27 RETURN node;
28 end;

```

Η παρακάτω συνάρτηση βρίσκει το πλησιέστερο κόμβο ως πηγή ή στόχος του πλησιέστερου συνδέσμου (link) εντός της δεδομένης απόστασης. Έχει τρία ορίσματα: το point, το distance όπου η συνάρτηση θα αναζητήσει για ένα κόμβο εντός αυτής της απόστασης, table όπου είναι το όνομα του πίνακα με τα δεδομένα. Η συνάρτηση αυτή αρχικά ψάχνει για το πλησιέστερο link. Αν δεν βρεθεί επιστρέφει null, διαφορετικά αν βρέθηκε το link γίνεται έλεγχος για το τι είναι πιο κοντά του το source ή το target και τέλος επιστρέφει το id του και την λέξη source ή target ανάλογα με το αποτέλεσμα του έλεγχου.

```

1  Function Εύρεση_κόμβου_από_κοντινότερο_link_εντός_απόστασης (point, distance, table)
2  Begin
3      -- Ψάχνοντας για το πλησιέστερο link
4      For row in 'select id from Εύρεση_κοντινότερου_link_εντός_απόστασης (point, distance, table) as id'
5      End Loop;
6      If row.id is null THEN
7          res.id = -1;
8          Return res;
9      End If;
10     link := row.id;
11     -- Έλεγχος τι είναι πιο κοντά - source ή target
12     For row in 'select distance ((select StartPoint (geom) from table where gid=link), point) as dist'
13     End Loop;
14     d1:= row.dist;
15     For row in 'select distance ((select EndPoint (geom) from table where gid=link), point) as dist'
16     End Loop;
17     d2:= row.dist;
18     IF d1<d2 THEN
19         field:= 'source';
20     ELSE
21         field:= 'target';
22     END IF;
23     res.id:= row.id;
24     res.name:= field;
25     Return res;
26 end;

```

Η παρακάτω συνάρτηση ταιριάζει μια δεδομένη γραμμή με το υπάρχον δίκτυο και επιστρέφει ένα σύνολο ακμών. Έχει έξι ορίσματα: το table όπου είναι το όνομα του πίνακα με τα δεδομένα, το line είναι η γραμμή όπου η συνάρτηση θα ταιριάζει με το υπάρχον δίκτυο, το dist είναι η απόσταση για την αναζήτηση του κοντινότερου κόμβου, το dist2 είναι η απόσταση για την αναζήτηση του συντομότερου μονοπατιού, το dir είναι true αν το γράφημα του δικτύου είναι κατευθυνόμενο, το rc είναι true αν υπάρχει reverse cost. Η συνάρτηση αυτή αρχικά ψάχνει για όλα τα σημεία της δεδομένης γραμμής. Στη συνέχεια μέσω μιας επανάληψης ψάχνει για τον πλησιέστερο κόμβο στο τρέχον σημείο εντός της απόστασης dist. Αν δεν υπάρχει πλησιέστερος κόμβος εντός δεδομένης απόστασης, δοκιμάζει τον αλγόριθμο για την εύρεση κόμβου από κοντινότερο link εντός απόστασης. Η επανάληψη συνεχίζεται για όλα τα σημεία της δεδομένης γραμμής και η συνάρτηση επιστρέφει το id των ακμών, το id των κόμβων και το κόστος της διαδρομής, διαφορετικά δεν μπορεί να βρει διαδρομή αν είναι null.

```

1  Function Αντιστοίχιση_γραμμής (table, line, dist, dist2, dir, rc)
2  Begin
3  num: = NumPoints (line);
4  -- Ψάχνοντας μέσα από όλα τα σημεία σε δεδομένη γραμμή
5  LOOP
6  i:= i+1;
7  -- Λήψη πλησιέστερου κόμβου στο τρέχον σημείο
8  For row in 'select * from Εύρεση_κοντινότερου_κόμβου_εντός_απόστασης (Point (line, i), dist, table) as id'
9  End Loop;
10 IF row.id IS NOT NULL THEN
11     points [i-1]:=row.id;
12 ELSE
13     -- Αν δεν υπάρχει πλησιέστερος κόμβος εντός δεδομένης απόστασης, δοκιμασία άλλου αλγόριθμου
14     For row in 'select * from Εύρεση_κόμβου_από_κοντινότερο_link (Point (line, i), dist2, table) as id'
15     End Loop;
16     points [i-1]:=row.id;
17     IF row.id = -1 THEN
18         return;
19     END IF;
20 END IF;
21 IF i>1 AND points [i-2] <> points [i-1] THEN
22     -- Κατασκευάζοντας το κύριο ερώτημα
23     query := 'select gid, geom from shortest_path (select gid, source, target, cost, x1, x2, y1, y2';
24     IF rc THEN query: = query ', reverse_cost';
25     END IF;
26     query: = query 'from table where (('Point(line,i-1)) - dist2*2', 'Point(line,i) + dist2*2') && geom,
27                                     'points[i-1]', 'points[i-2]', dir, rc, table), where edge_id=gid';
28 FOR row IN EXECUTE query LOOP
29     IF row IS NULL THEN
30         'Δεν μπορεί να βρει μια διαδρομή μεταξύ ', points[i-1], points[i-2];
31         RETURN;
32     END IF;
33     edges[z] := row.edge_id; vertices[z] := row.vertex_id; costs[z] := row.cost;
34     IF edges[z] = -1 THEN t := 0;

```

```

37     -- Ordering edges
38     FOR t IN (prev+1)..z-1 LOOP
39         path.edge_id := edges[t]; path.vertex_id := vertices[t]; path.cost = costs[t];
40         edges[t] := edges[z-t+prev+1]; vertices[t] := vertices[z-t+prev+1]; costs[t] := costs[z-t+prev+1];
41         edges[z-t+prev+1] := path.edge_id; vertices[z-t+prev+1] := path.vertex_id; costs[z-t+prev+1] := path.cost;
42     END LOOP;
43     prev := z;
44 END IF;
45 z := z+1;
46 END LOOP;
50 END IF;
51 EXIT WHEN i=num;
52 END LOOP;
53 FOR t IN 0..array_upper(edges, 1) LOOP
54     RETURN NEXT path;
55 END LOOP;
62 END;

```

Η παρακάτω συνάρτηση ταιριάζει μια δεδομένη γραμμή με το υπάρχον δίκτυο και επιστρέφει ένα linestring ή multilinestring. Έχει έξι ορίσματα: το table όπου είναι το όνομα του πίνακα με τα δεδομένα, το line είναι η γραμμή όπου η συνάρτηση θα ταιριάζει με το υπάρχον δίκτυο, το dist είναι η απόσταση για την αναζήτηση του κοντινότερου κόμβου, το dist2 είναι η απόσταση για την αναζήτηση του συντομότερου μονοπατιού, το dir είναι true αν το γράφημα του δικτύου είναι κατευθυνόμενο, το rc είναι true αν υπάρχει στήλη reverse cost. Η συνάρτηση αυτή αρχικά καλεί την συνάρτηση Αντιστοίχιση_γραμμής. Αν το αποτέλεσμα από την συνάρτηση αυτή είναι μηδέν επιστρέφει null, διαφορετικά προσπαθεί να δημιουργήσει μια linestring ή μια multilinestring. Ελέγχει αν το αποτέλεσμα είναι έγκυρο και επιστρέφει μια γεωμετρία δηλαδή την linestring ή την multilinestring αλλιώς ένα μήνυμα ότι το αποτέλεσμα δεν είναι έγκυρη γεωμετρία.

```

1  Function Αντιστοίχιση_γραμμής_ως_linestring (table, line, dist, dist, dir, rc)
2  Begin
3  i:= 0;
4  For row in 'select * from Αντιστοίχιση_γραμμής (table, line, dist, dist, dir, rc)' Loop
5      edges[i] := row.edge_id;
6      i:= i+1;
7  End Loop;
8  IF i = 0 THEN
9      return NULL;
10 END IF;
11 -- Προσπάθεια να δημιουργηθεί μια linestring. Επίσης μπορεί να επιστρέψει multilinestring.
12 For row in 'select linemerge (geomunion (multi (geom))) as geom from table where gid > 0' Loop
13 End Loop;
14 IF isvalid(row.geom) THEN
15     RETURN row.geom;
16 ELSE
17     'Το αποτέλεσμα δεν είναι έγκυρη γεωμετρία.';
18 END IF;
19 end;

```

4.4 Περιγραφή Υλοποίησης του Map Matching

Στην αρχή, γίνεται η εγκατάσταση της PostgreSQL. Στη συνέχεια για να παρέχει η PostgreSQL υποστήριξη σε χωρικά δεδομένα και δρομολόγηση γίνεται η εγκατάσταση του PostGIS και του PgRouting. Μετά δημιουργείται η βάση δεδομένων όπου θα αποθηκευτούν τα χωρικά δεδομένα. Για της ανάγκες τις εργασίας δημιουργήθηκαν τρεις πίνακες οι οποίοι είναι οι παρακάτω:

Ο πίνακας `milano_network` περιέχει τα δεδομένα του οδικού δικτύου της πόλης του Μιλάνου. Τα δεδομένα του οδικού δικτύου είναι σε μορφή `osm` και μέσω της εφαρμογής `osm2po` μετατρέπονται σε `sql` αρχείο και στη συνέχεια μπορούν να εισαχθούν στη βάση δεδομένων. Το `Osm2po` είναι και μετατροπέας και μια μηχανή δρομολόγησης. Αναλύει τα δεδομένα `OpenStreetMap` και καθιστά δυνατά να δρομολογηθούν. Παράγει `SQL` δεδομένα για `PostGIS`, συμβατά με `pgRouting` και `Quantum GIS`.

Αρχικά, γίνεται λήψη του `OSM` δυαδικού αρχείου, `milano.osm.pbf`. Η φόρτωση του δικτύου θα γίνει από την απευθείας εκτέλεση από την γραμμή εντολών μέσω της παρακάτω εντολής:

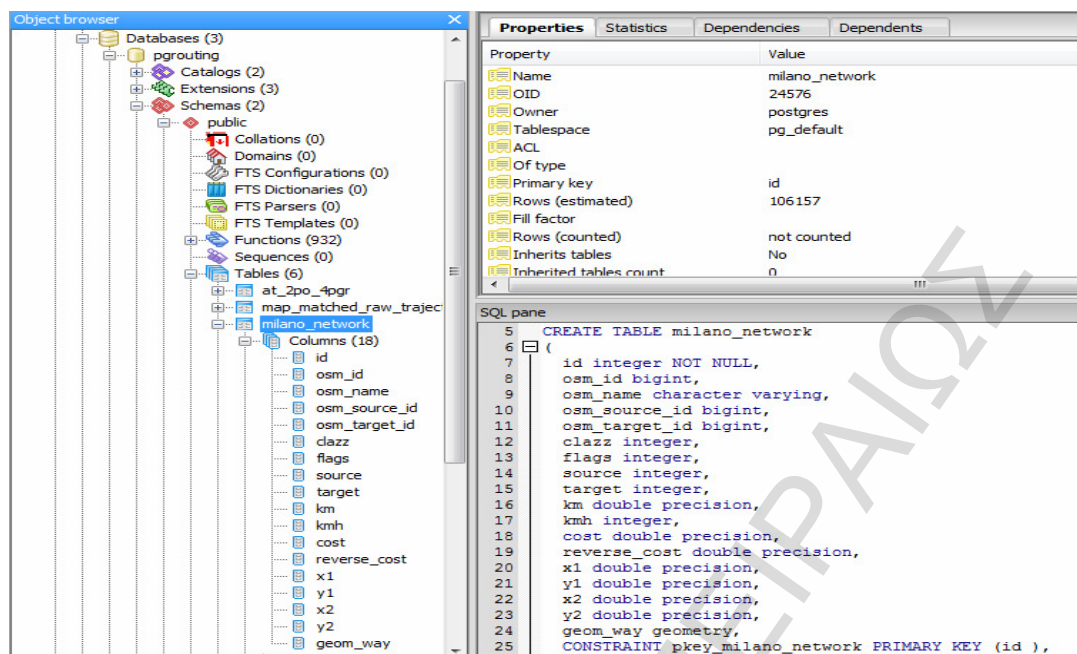
```
C:\Users\tina\temp\osm2po-4.2.30> java -jar osm2po-core-4.2.30-signed.jar
prefix=at "C:\Users\tina\Geodaten\OpenStreetMap Data\milano_network.osm.pbf"
```

Τώρα μπορεί εύκολα να εισάγεται το αρχείο `.sql` σε μια υπάρχουσα βάση δεδομένων. Στην περίπτωση αυτή η ενεργοποιημένη `PgRouting` βάση δεδομένων είναι η `pgrouting`.

```
C:\Users\tina\temp\osm2po-4.2.30\at>psql -U [username] -d pgrouting -q -f
C:\Users\tina\temp\osm2po-4.2.30\at\milano_network.sql
```

Τώρα, τα δεδομένα είναι έτοιμα για χρήση. Μετά τις παραπάνω εντολές πραγματοποιείται η εισαγωγή των δεδομένων και εμφανίζεται ο πίνακας `milano_network` στη βάση δεδομένων `pgrouting`. Τα πεδία του πίνακα φαίνονται παρακάτω: Το `id` είναι το πρωτεύων κλειδί του πίνακα, άλλα σημαντικά πεδία είναι το `source`, το `target`, το `cost`, το `reverse_cost`, τα σημεία `x` και `y` και την γεωμετρία `geom_way`.

```
CREATE TABLE milano_network
(
  id integer NOT NULL,
  osm_id bigint,
  osm_name character varying,
  osm_source_id bigint,
  osm_target_id bigint,
  clazz integer,
  flags integer,
  source integer,
  target integer,
  km double precision,
  kmh integer,
  cost double precision,
  reverse_cost double precision,
  x1 double precision,
  y1 double precision,
  x2 double precision,
  y2 double precision,
  geom_way geometry,
  CONSTRAINT pkey_milano_network PRIMARY KEY (id)
);
```



Εικόνα 4-1: Ο πίνακας του δικτύου milano_network

Ο δεύτερος πίνακας που δημιουργήθηκε είναι ο milano_seg_pgis_20percent στον οποίο θα εισαχθούν τα χωροχρονικά δεδομένα. Τα χωροχρονικά δεδομένα βρίσκονται σε ένα txt αρχείο, με τιμές για τα κάτωθι μεγέθη:

| Πεδίο | Περιγραφή | Τύπος |
|----------|---|-----------------------------|
| obj_id | Το id του κινουμένου αντικειμένου | integer |
| traj_id | Το id της διαδρομής | integer |
| seg_id | Το id του segment | integer |
| seg_pgis | Μια linestring γεωμετρία | geometry |
| ti | Η χρονική στιγμή έναρξης για ένα συγκεκριμένο segment | timestamp without time zone |
| te | Η χρονική στιγμή τέλους στο συγκεκριμένο segment | timestamp without time zone |

Πίνακας 4-1: Περιγραφή Χωροχρονικών Δεδομένων

Έγιναν οι πιο κάτω διαδικασίες:

Δημιουργήθηκε ο milano_seg_pgis_20percent πίνακας για τα χωροχρονικά δεδομένα που δίνονται, στη Σχισιακή Βάση Δεδομένων Postgresql. Έχει ένα σύνθετο πρωτεύων κλειδί που αποτελείται από τρία πεδία τα obj_id, traj_id, seg_id.

Δόθηκε η SQL εντολή:

```
CREATE TABLE milano_seg_pgis_20percent
(
  obj_id integer NOT NULL,
  traj_id integer NOT NULL,
```

```

seg_id integer NOT NULL,
seg_pgis geometry,
ti timestamp without time zone,
te timestamp without time zone,
CONSTRAINT milano_seg_pgis_20percent_pkey PRIMARY KEY (obj_id, traj_id, seg_id)
);

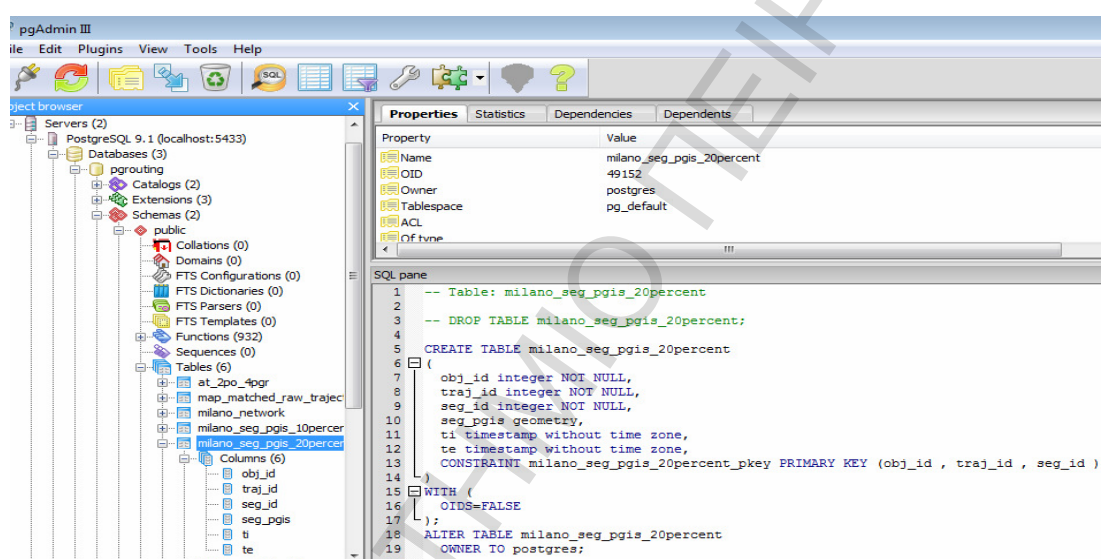
```

Και στη συνέχεια για το φόρτωμα των δεδομένων από το αρχείο C:\Program Files\PostgreSQL\9.1\share\milan_subset_2_seg_pgis.txt στον πίνακα milano_seg_pgis_20percent που δημιουργήθηκε παραπάνω, δόθηκε η παρακάτω SQL εντολή:

```

COPY milano_seg_pgis_20percent
FROM 'C:\Program Files\PostgreSQL\9.1\share\milan_subset_2_seg_pgis.txt';

```



Εικόνα 4-2: Ο πίνακας με τα χωροχρονικά δεδομένα

Τέλος, ο τρίτος πίνακας που δημιουργήθηκε είναι ο map_matched_raw_trajectory στον οποίον θα εισαχθούν τα αποτελέσματα του map-matching. Το output της συνάρτησης Αντιστοίχιση_γραμμής_ως_linestring θα εισαχθεί σε αυτόν τον πίνακα.

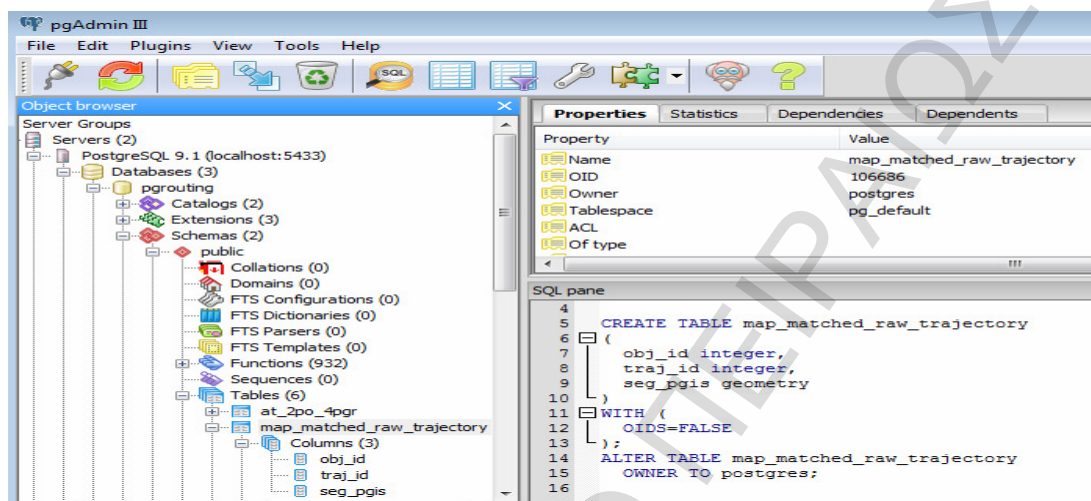
| Πεδίο | Περιγραφή | Τύπος |
|----------|-----------------------------------|----------|
| obj_id | Το id του κινουμένου αντικειμένου | integer |
| traj_id | Το id της διαδρομής | integer |
| seg_pgis | Μια linestring γεωμετρία | geometry |

Πίνακας 4-2: Ο πίνακας με τα map-matched αποτελέσματα

Έγιναν οι πιο κάτω διαδικασίες :

Δημιουργήθηκε ο map_matched_raw_trajectory πίνακας για τα map-matching αποτελέσματα. Δόθηκε η SQL εντολή:


```
CREATE TABLE map_matched_raw_trajectory
(
  obj_id integer,
  traj_id integer,
  seg_pgis geometry
);
```



Εικόνα 4-3: Ο πίνακας με τα map-matched αποτελέσματα

| Συναρτήσεις του PostGIS που χρησιμοποιήθηκαν |
|--|
| ST_Distance |
| ST_PointN |
| ST_NPoints |
| ST_X |
| ST_Y |
| ST_StartPoint |
| ST_EndPoint |
| ST_GeometryType |
| ST_GeometryFromType |
| ST_SetSrid |
| ST_Srid |
| ST_Astext |
| ST_linemerge |
| ST_Union |
| ST_Multi |
| ST_Isvalid |

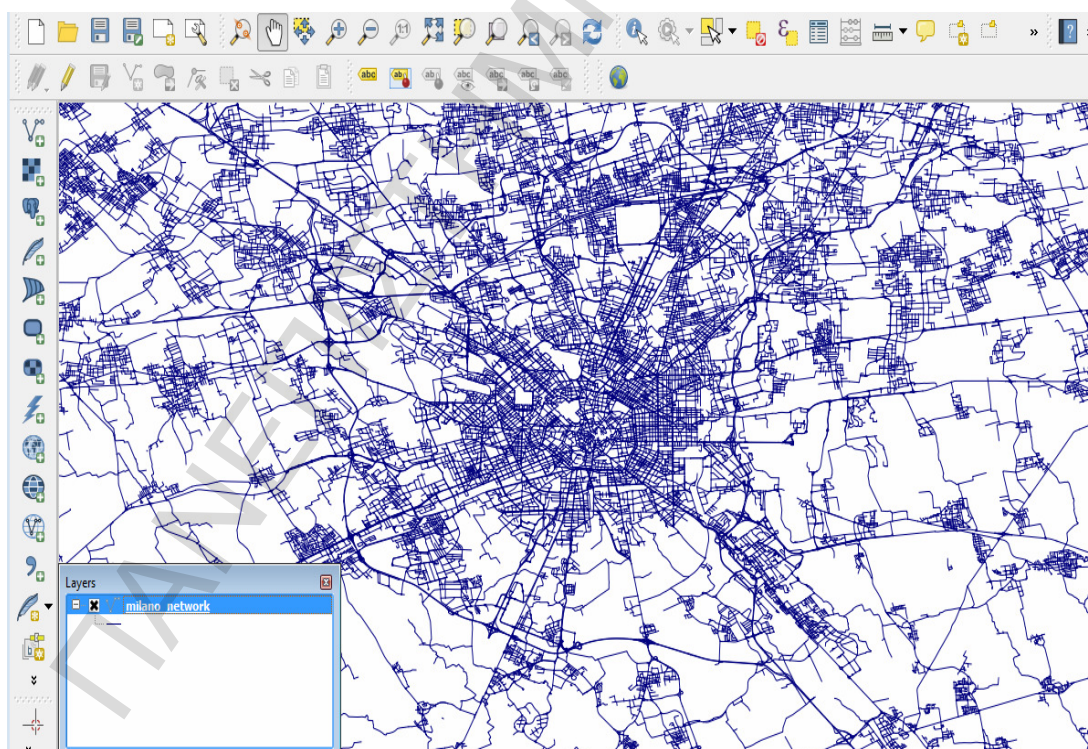
Πίνακας 4-3: Συναρτήσεις του PostGIS που χρησιμοποιήθηκαν

4.5 Οπτικοποίηση Αποτελεσμάτων

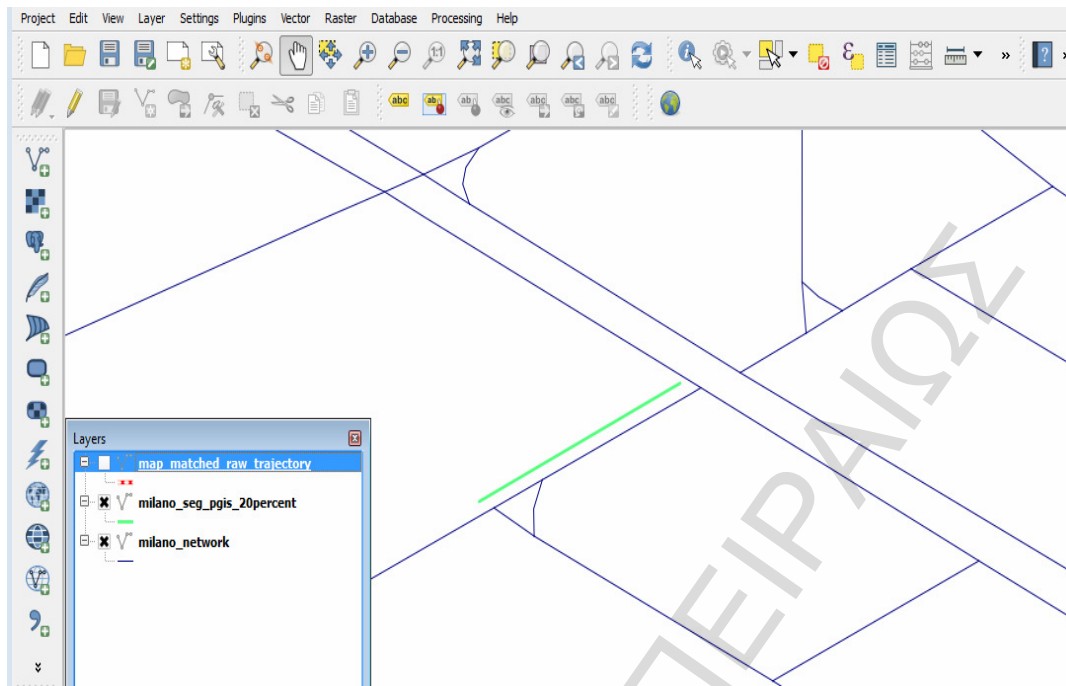
Τα αποτελέσματα έχουν οπτικοποιηθεί στο QGIS. Το QGIS είναι ένα ελεύθερο και ανοιχτού κώδικα Σύστημα Γεωγραφικών Πληροφοριών (GIS) που παρέχει προβολή δεδομένων, επεξεργασία και ανάλυση δυνατοτήτων. Παρόμοια με άλλα GIS συστήματα, το λογισμικό QGIS επιτρέπει στους χρήστες να δημιουργήσουν χάρτες με πολλά επίπεδα που χρησιμοποιούν διαφορετικές χαρτογραφικές προβολές.

Η διαδικασία που γίνεται είναι η εξής: Το QGIS δίνει την δυνατότητα σύνδεσης με την PostgreSQL. Μέσω της επιλογής “Add PostGIS Layers” δημιουργούμε στην αρχή μια καινούργια σύνδεση δηλαδή μια “Create a new PostGIS connection” δίνοντας όλες τις πληροφορίες της σύνδεσης (Name, Service, Host, Port, Database, SSL Mode, Username, Password). Αφού δημιουργήθηκε η PostGIS σύνδεση, μετά επιλέγουμε το name που δόθηκε πριν και πατώντας το κουμπί «connect» γίνεται η σύνδεση και μπορούμε να προσθέσουμε το πίνακα της βάσης δεδομένων που μας ενδιαφέρει. Αφού επιλέξουμε το πίνακα και πατώντας το κουμπί «add» βλέπουμε στο παράθυρο με τα layers να έχει γίνει η πρόσθεση ενός καινούργιου επιπέδου με το όνομα του πίνακα που επιλέξαμε.

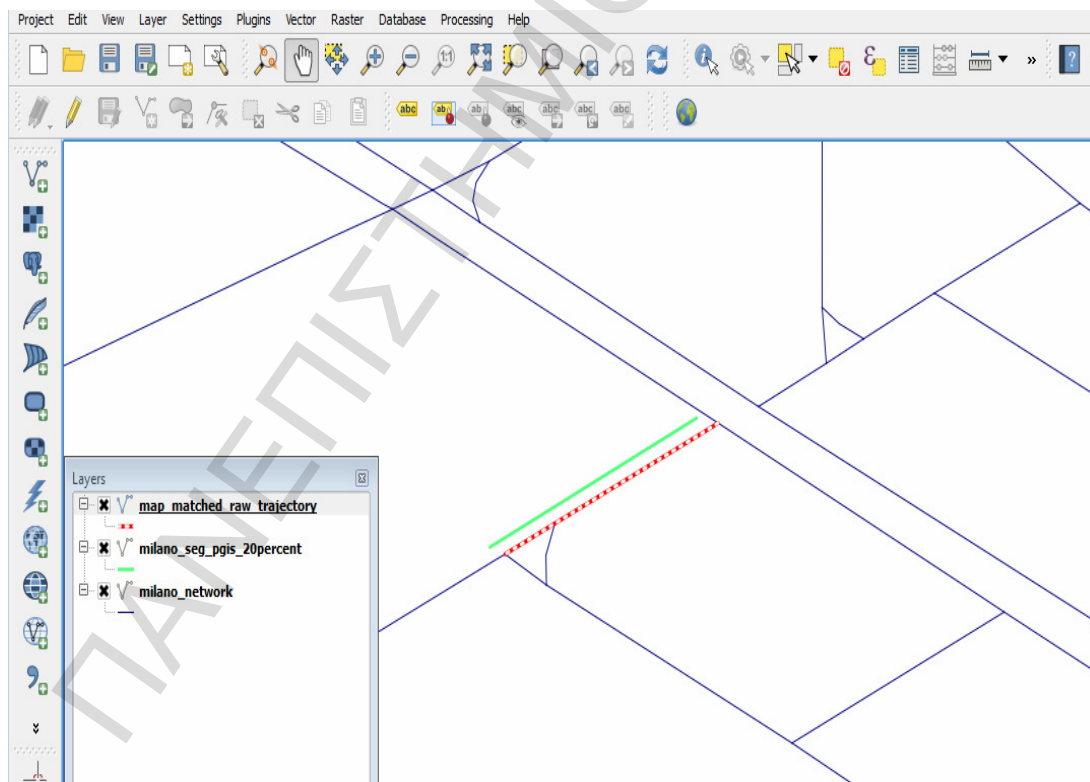
Όπως φαίνεται στις εικόνες η παραπάνω διαδικασία έχει γίνει τρεις φορές, αφού έχουμε τρία layers (επίπεδα). Το ένα επίπεδο είναι το Milano_network όπου απεικονίζεται το δίκτυο της πόλης του Μιλάνου. Το άλλο επίπεδο είναι το Milano_seg_rgis_20percent όπου απεικονίζονται τα δεδομένα πάνω στα οποία θα γίνει το map_matching δηλαδή το dataset. Το τελευταίο επίπεδο είναι το map_matched_raw_trajectory όπου απεικονίζονται τα map-matching αποτελέσματα. Άρα, όπως φαίνεται στις παρακάτω εικόνες οι μπλε γραμμές είναι οι δρόμοι του οδικού δικτύου του Μιλάνου. Οι πράσινες γραμμές είναι οι εκτιμώμενες θέσεις από τις οποίες έχει περάσει το κινούμενο αντικείμενο. Οι κόκκινες γραμμές είναι το map-matching αποτέλεσμα δηλαδή οι πραγματικές θέσεις από τις οποίες έχει περάσει το κινούμενο αντικείμενο. Παρακάτω βλέπουμε μερικά παραδείγματα και τα map-matching αποτελέσματα από κάποιες τροχιές.



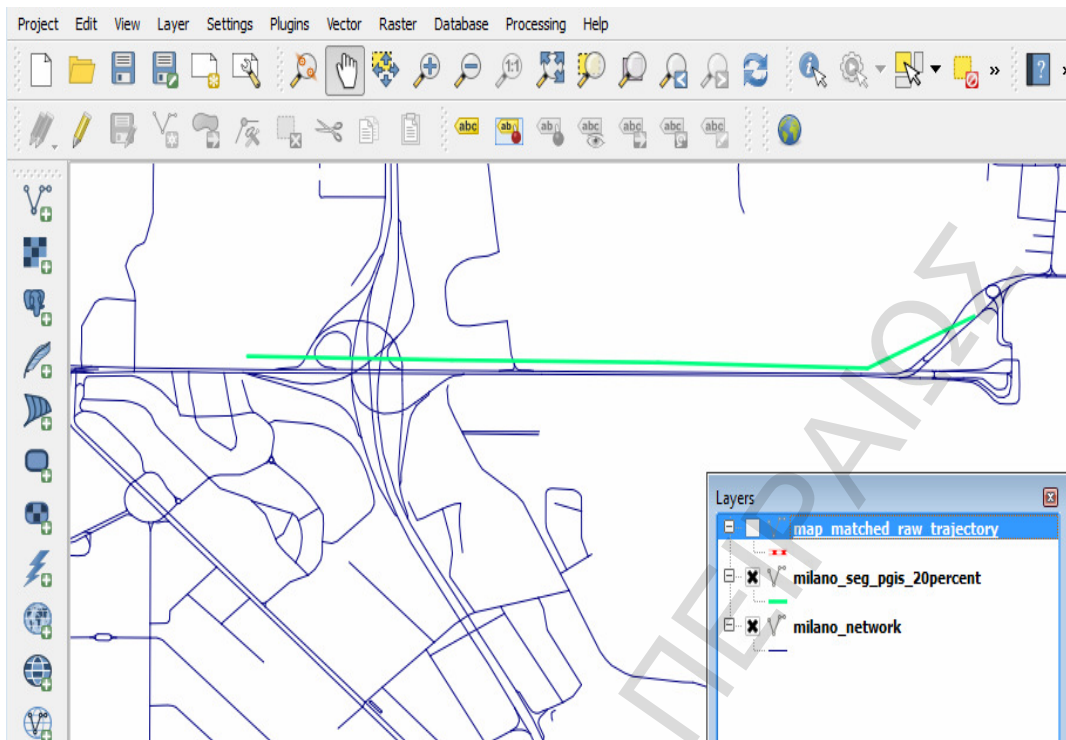
Εικόνα 4-4: Ο πίνακας του δικτύου του Μιλάνου



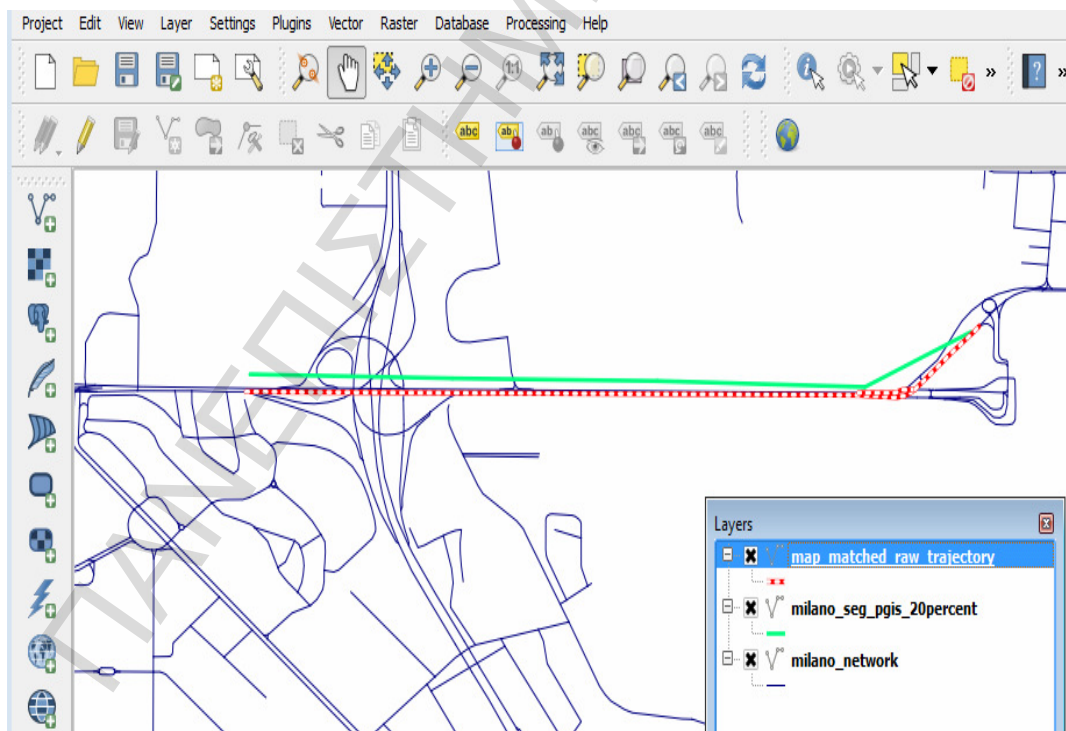
Εικόνα 4-5: Παράδειγμα 1



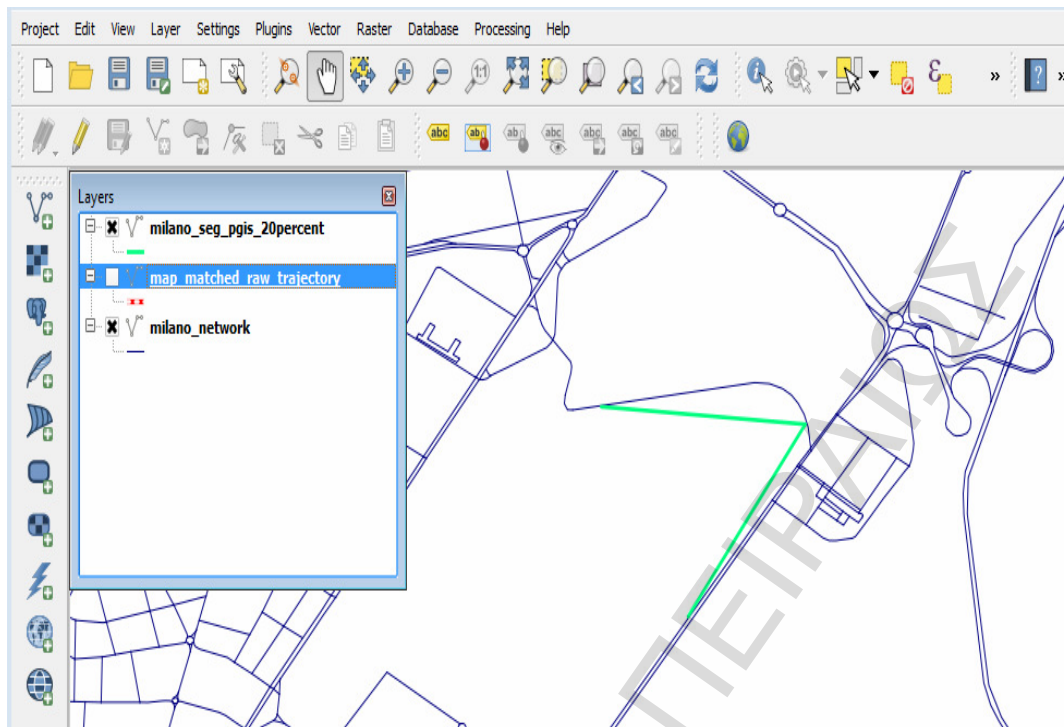
Εικόνα 4-6: Το map-matching αποτέλεσμα του παραδείγματος 1



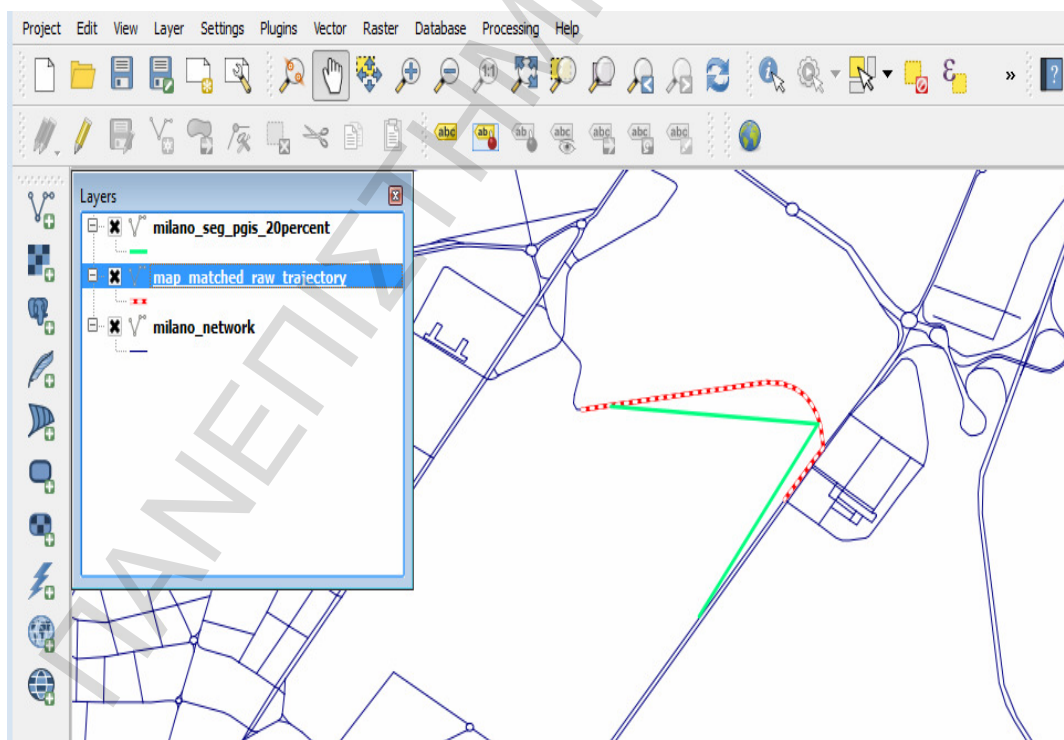
Εικόνα 4-7: Παράδειγμα 2



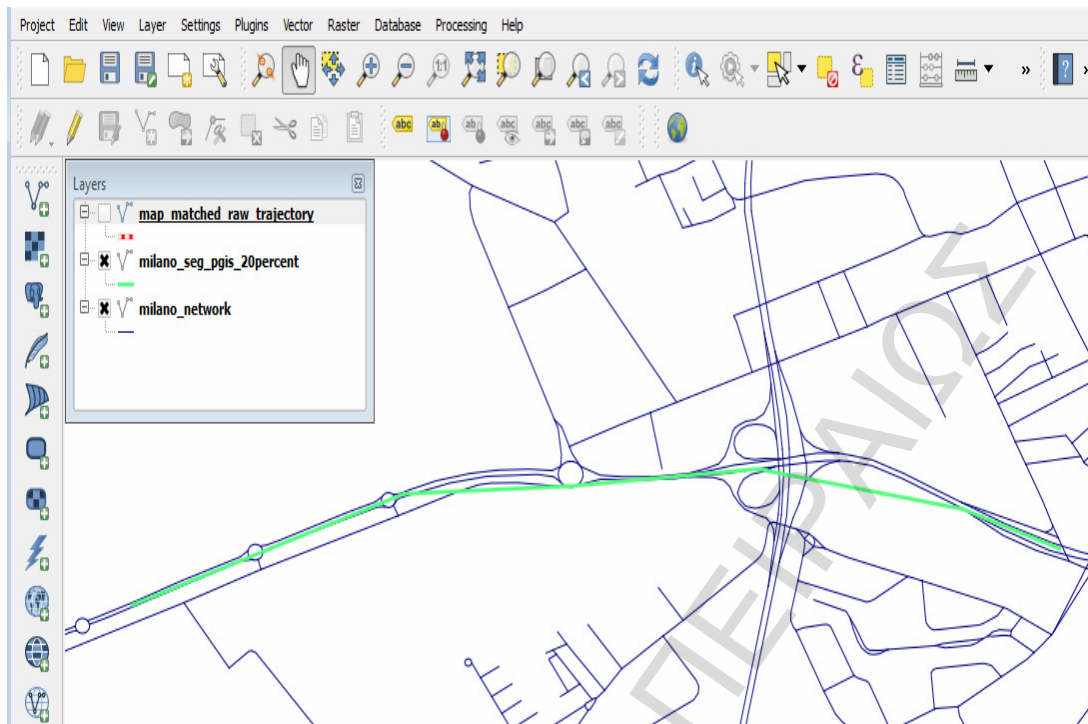
Εικόνα 4-8: Το map-matching αποτέλεσμα του παραδείγματος 2



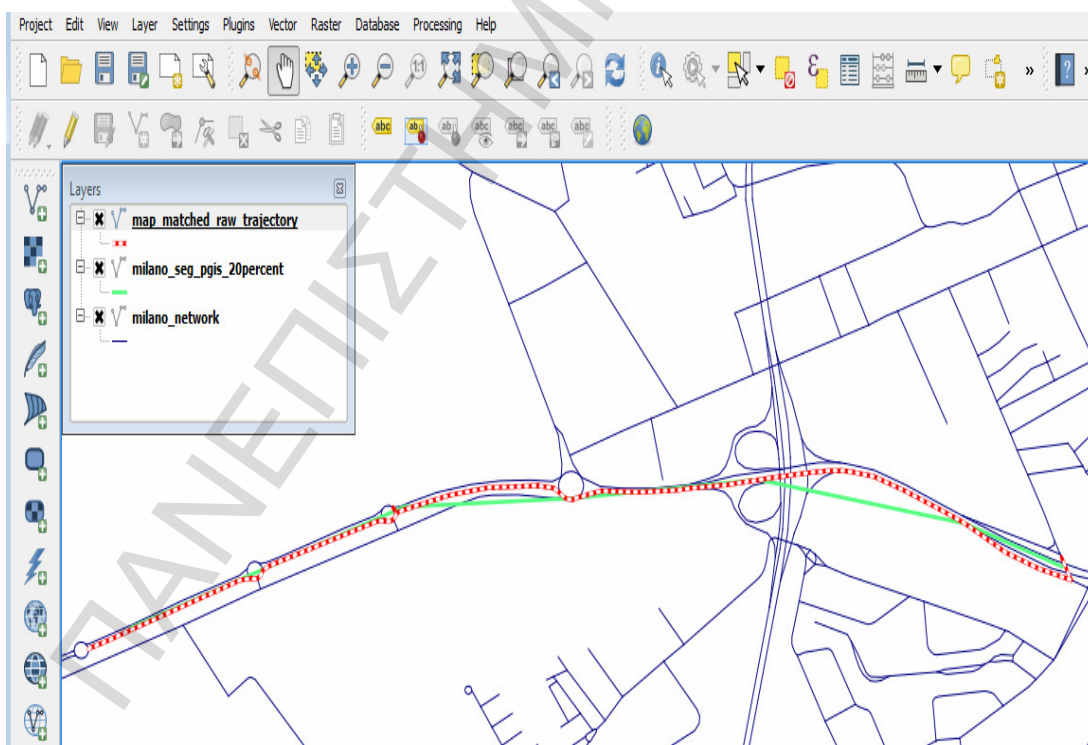
Εικόνα 4-9: Παράδειγμα 3



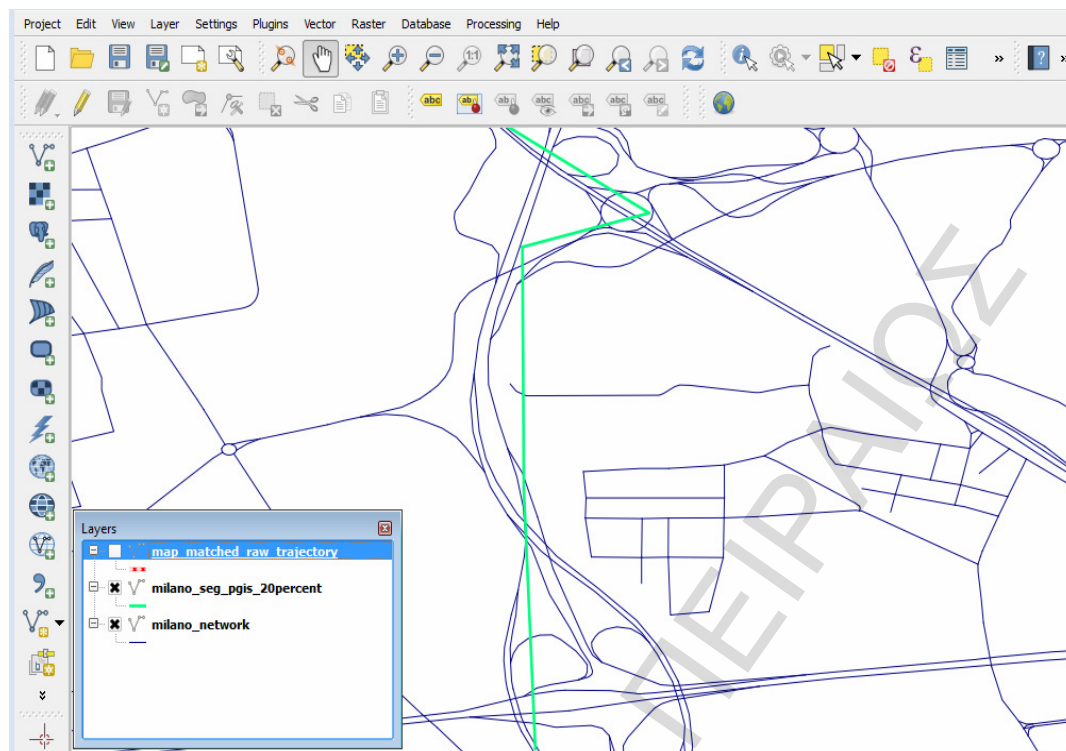
Εικόνα 4-10: Το map-matching αποτέλεσμα του παραδείγματος 3



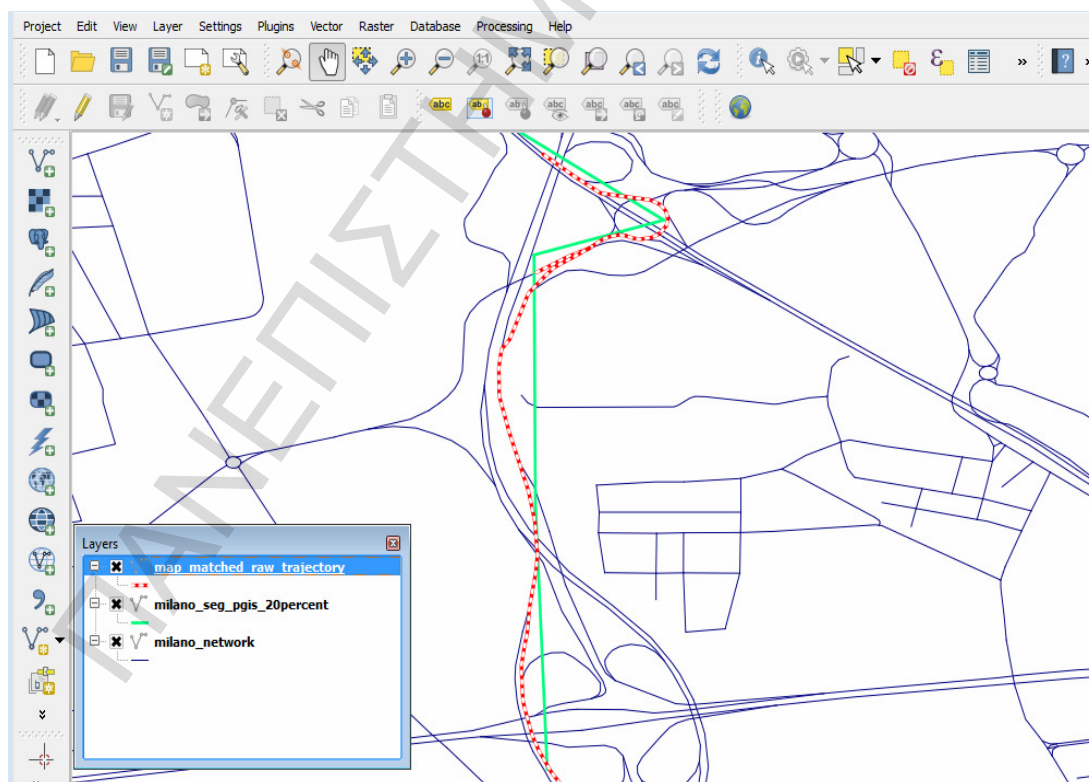
Εικόνα 4-11: Παράδειγμα 4



Εικόνα 4-12: Το map-matching αποτέλεσμα του παραδείγματος 4



Εικόνα 4-13: Παράδειγμα 5



Εικόνα 4-14: Το map-matching αποτέλεσμα του παραδείγματος 5

ΣΥΜΠΕΡΑΣΜΑΤΑ

Ολοκληρώνοντας την παρούσα διπλωματική εργασία πραγματοποιήθηκε με επιτυχία η εύρεση της πιθανής περιοχής δραστηριοποίησης ενός κινούμενου αντικείμενου, εστιάζοντας στην χώρο-χρονική μελέτη της συνεχούς κίνησης ανθρώπου (ή γενικότερα κινούμενου αντικείμενου) μεταξύ δύο σημείων, που θεωρούνται αφετηρία και τέλος διαδρομής, κινείται με μεταβλητή ταχύτητα, είναι γνωστή δε η μέγιστη ταχύτητα που μπορεί να επιτύχει. Η δημιουργία της συνάρτησης (PAA) εύρεσης της πιθανής περιοχής δραστηριοποίησης ενός κινούμενου αντικείμενου και τα αποτελέσματα της συνάρτησης αυτής επιβεβαιώνουν το αποτέλεσμα.

Επίσης, το δεύτερο ζήτημα που μελετήθηκε και υλοποιήθηκε στα πλαίσια της εργασίας είναι το map-matching. Η υλοποίηση του συγκεκριμένου προβλήματος ολοκληρώθηκε με επιτυχία όπως φαίνεται και στα αποτελέσματα που παρουσιάστηκαν.

Τέλος, μία μελλοντική εργασία είναι να χρησιμοποιηθεί μία άλλη προσέγγιση αυτού του προβλήματος, δηλαδή να προσπαθήσουμε να βρούμε την πιθανή περιοχή δραστηριοποίησης του κινούμενου αντικείμενου σε ένα δίκτυο. Λαμβάνοντας υπόψη ένα οδικό δίκτυο, ένα κινούμενο αντικείμενο O , n timestamps t_i , $i = 1 \dots m \dots n$ (η διαδοχικά δείγματα), n χωρικές θέσεις x_i , $i = 1 \dots m \dots n$ (οι θέσεις του αντικείμενου στο t_i) και η μέγιστη ταχύτητα M που μπορεί να επιτευχθεί από το συγκεκριμένο κινούμενο αντικείμενο.

Μία άλλη μελλοντική εργασία είναι να χρησιμοποιηθεί η συσχέτιση του PAA με το map-matching και το δίκτυο που να δείχνει την περιοχή που είναι κοινή για ένα σύνολο από PAAs, διαφορετικών εν κινήσει χρηστών. Δηλαδή, μπορούμε να βρούμε και να προτείνουμε σημεία συνάντησης ως τα κοινά σημεία ενδιαφέροντος που υπάρχουν μέσα σε αυτή την περιοχή.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Βιβλιογραφικές Αναφορές

- [PJ1999] Dieter Pfoser and Christian S. Jensen (1999): Capturing the Uncertainty of Moving-Object Representations pp. 111-131.
- [NVS2010] Tijs Neutens, Mathias Versichele, Tim Schwanen (2010): Arranging place and time: A GIS toolkit to assess person-based accessibility of urban opportunities pp. 561-575.
- [NWD2007] Tijs Neutens, Frank Witlox, Philippe Demaeyer (2007). Individual accessibility and travel possibilities: A literature review on time geography, pp. 335-352.
- [NWM2007] Neutens, T., Witlox, F., Van de Weghe, N., & De Maeyer, P. (2007). Space-time opportunities for multiple agents: a constraint-based approach. *International Journal of Geographical Information Science*, 21(10), 1061–1076.
- [AR1999] Abraham, Tamas; Roddick, John F.; "Survey of Spatio-Temporal Databases", *Geoinformatica*, pp. 61-99, vol. 3, 1, 1999
- [GBEJ2000] Güting, R. H., M. H. Böhlen, M. Erwig, C. S. Jensen, N. A. Lorentzos, M. Schneider, M. Vazirgiannis, 2000. "A foundation for representing and querying moving objects", pp. 1-42.
- [CSZ2011] Z. Chen, H. T. Shen, and X. Zhou, "Discovering Popular Routes from Trajectories," *Proc. IEEE ICDE*, pp. 900-911, 2011.
- [ZZX2012] K. Zheng, Y. Zheng, X. Xie, and X. Zhou, "Reducing Uncertainty of Low-Sampling-Rate trajectories," *Proc. IEEE ICDE*, pp. 1144- 1155, 2012.
- [BK1996] D. Bernstein, A. Kornhauser, "An introduction to map matching for personal navigation assistants". New Jersey TIDE Center, 1996.
- [WBK2000] C. E. White, D. Bernstein and A. L. Kornhauser, "Some Map Matching Algorithms for Personal Navigation System", *Transportation Research Part C*, Vol. 8, 2000.
- [ZG2006] J. Zhou, R. Golledge, 2006 "A Three-step General Map Matching Method in the GIS Environment: Travel/Transportation Study Perspective". *International Journal of Geographical Information System*.
- [AWB2011] Ahmedin M. Ahmed, Weihua Zhu, Teshome M. Bekele. Map-matching and Positioning Uncertainty in Location Based Services (LBS). *International Conference on Asia Agriculture and Animal IPCBEE* vol.13 (2011).
- [LZZX2009] Y Lou, C Zhang, Y Zheng, X Xie, W Wang, Y Huang, 2009. Map-Matching for Low-Sampling-Rate GPS Trajectories. *Proceedings of ACM SIGSPATIAL Conference on Geographical Information Systems (ACM GIS 2009)*.

[ΣΕ2003] Στεφανάκης Εμμανουήλ (2003), Βάσεις Γεωγραφικών Δεδομένων και Συστήματα Γεωγραφικών Πληροφοριών, Εκδόσεις Παπασωτηρίου, Αθήνα.

[ΘΠ2012] Γ. Θεοδωρίδης, Ν. Πελέκης (2012), Διδακτικές Σημειώσεις Μαθήματος: Γεωγραφικά Πληροφορικά Συστήματα, ΠΜΣ – Προηγμένα Συστήματα Πληροφορικής, Πειραιάς.

[BT2012] Εργαστηριακή Διάλεξη στα Γεωγραφικά Πληροφορικά Συστήματα – PostGIS - Μάριος Βόντας, Παναγιώτης Ταμπάκης 2012.

[PSQL] PostgreSQL User Guide, Release 9.2, The PostgreSQL Global Development Group.

ΔΙΚΤΥΑΚΟΙ ΤΥΠΟΙ

[URL1] Χωροχρονικές Βάσεις Δεδομένων:

<http://www-users.cs.umn.edu/~lmani/spatial/HW5-B4.pdf>

[URL2] Χωροχρονικές Βάσεις Δεδομένων:

http://en.wikipedia.org/wiki/Spatiotemporal_database

[URL3] Κινούμενα αντικείμενα:

http://www.dbnet.ece.ntua.gr/~kpatro/MSc_Dissertation/Text/Chapter4.pdf

[URL4] Διδακτορική Διατριβή Η. Φρέντζος:

<http://digilib.lib.unipi.gr/dspace/bitstream/unipi/2996/1/Frentzos.pdf>

[URL5] Space Time Accessibility Toolkit:

http://cartogis.ugent.be/sta_toolkit

[URL6] PostgreSQL Site:

<http://www.postgresql.org/>

[URL7] PostGIS Site:

<http://postgis.net/>

[URL8] PostGIS 2.0 Manual:

<http://postgis.net/docs/manual-2.0/>

[URL9] pgRouting Site:

<http://pgrouting.org/>

[URL10] pgRouting Workshop Manual:

<http://workshop.pgrouting.org/>

[URL11] OSGeo Live:

http://live.osgeo.org/el/overview/pgrouting_overview.html

[URL12] PostgreSQL:

<http://el.wikipedia.org/wiki/PostgreSQL>

[URL13] PostGIS:

<http://infolab.cs.unipi.gr/pre-eclass/courses/gis/lab/Lab-PostGIS-ppt.pdf>

[URL14] pgRouting:
<http://docs.pgRouting.org/2.0/en/pgRoutingDocumentation.pdf>

[URL15] To osm2po:
<http://osm2po.de/>

[URL16] To Google Earth:
<https://developers.google.com/earth/documentation/>

[URL17] KML Documentation:
<https://developers.google.com/kml/documentation/>

[URL18] QGIS Site:
<http://www.qgis.org/en/site/>

[URL19] QGIS 2.0 Documentation:
http://www.qgis.org/en/docs/user_manual/index.html

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ