

ΑΝΑΖΗΤΗΣΗ ΕΓΓΥΤΕΡΩΝ ΓΕΙΤΟΝΩΝ ΣΕ ΑΒΕΒΑΙΑ ΣΤΙΓΜΑΤΑ ΑΠΟ ΚΙΝΗΤΑ

ΠΑΗΡΟΦΟΡΙΕΣ: Κώστας Πατρούμπας, 210 772 1446, kpatro@dblab.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Με την υποτιθέμενη υπηρεσία, ανά πάσα χρονική στιγμή ο χρήστης θα μπορεί να πληροφορείται ποιοί *κατά πάσα πιθανότητα* είναι οι k πλησιέστεροι φίλοι του με βάση προσεγγιστικά γεωγραφικά στίγματα από τα κινητά τους τηλέφωνα.

ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: C++/Java.

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Όσοι κάτοχοι κινητών συσκευών (smart phones) εγγράφονται συνδρομητές στην υπηρεσία, επιθυμούν να λαμβάνουν ειδοποιήσεις σχετικά με φίλους τους που βρίσκονται τώρα πλησιέστερα, λ.χ. για να κανονίσουν συνάντηση. Για λόγους προστασίας του απορρήτου, κανενός το γεωγραφικό στίγμα δεν δίνεται με ακριβείς συντεταγμένες, αλλά είναι επίτηδες *αβέβαιο*. Ωστόσο, πάντοτε θεωρείται γνωστή μία ευρύτερη περιοχή, όπως λ.χ. η χωρική κάλυψη της κυψέλης κινητής τηλεφωνίας ή μία εκτεταμένη ζώνη εκατοντάδων μέτρων όπου κάποιος εντοπίζεται με αρκετή, αλλά *κυμαινόμενη πιθανότητα*. Οι χρήστες θα μπορούν λοιπόν να υποβάλλουν *ερωτήματα διαρκείας* (continuous queries), ώστε όποιος θέλει, να ενημερώνεται σε πραγματικό χρόνο σχετικά με τους k *πιθανότερους εγγύτερους γείτονές του* (probabilistic k -nearest neighbors, όπου $k = 2, 3, 5, \dots$). Ο αλγόριθμος επεξεργασίας πρέπει ν' αντεπεξέρχεται σε μαζικές ανανεώσεις γεωγραφικών στιγμάτων με την μορφή *ρεύματος δεδομένων* (data stream) από χιλιάδες χρήστες, ενώ ταυτόχρονα οφείλει να δίνει τακτικά ενημερωμένες *προσεγγιστικές απαντήσεις*, αλλά με αποδεκτή ποιότητα.

Τα στάδια εκπόνησης της διπλωματικής εργασίας προβλέπονται ως εξής:

- Επισκόπηση μεθόδων επεξεργασίας πιθανοτικών δεδομένων & ερωτημάτων εγγύτερων γειτόνων.
- Σχεδιασμός του *πιθανοτικού μοντέλου* καταγραφής στιγμάτων και δήλωσης ερωτημάτων. Διάφορες παραλλαγές του ερωτήματος μπορούν να μελετηθούν, λ.χ. εντοπισμός γειτόνων με πιθανότητες πάνω από συγκεκριμένο κατώφλι ή πώς αποφασίζεται η σειρά κατάταξης των k γειτόνων σε καθεστώς *χωρικής αβεβαιότητας* (spatial uncertainty).
- Υλοποίηση του μηχανισμού αποτίμησης για *μία επιλεγμένη παραλλαγή* του ερωτήματος. Ο αλγόριθμος μάλλον θ' απαιτήσει χρήση ενός χωροχρονικού *ευρετηρίου* (index) προκειμένου να διευκολύνονται οι χωρικές αναζητήσεις.
- Αξιολόγηση πειραματικών αποτελεσμάτων σε δοκιμαστικά σύνολα δεδομένων, μετρώντας επιδόσεις για χρόνο απόκρισης, απαιτήσεις σε μνήμη, ακρίβεια απαντήσεων κ.ά.

Επισημαίνεται ότι η εργασία επικεντρώνεται αποκλειστικά στο ζήτημα της διαχείρισης δεδομένων κίνησης με σκοπό την χρήση τους στην αποτίμηση χωρικών ερωτημάτων και δεν πρόκειται να υπεισέλθει σε θέματα συλλογής, μετάδοσης ή αξιοπιστίας των στοιχείων. Οι ενδιαφερόμενοι θα πρέπει να είναι καλά καταρτισμένοι στα γνωστικά πεδία των βάσεων δεδομένων και της πιθανοθεωρίας και διατεθειμένοι να μελετήσουν ενδιαφέροντα γεωμετρικά προβλήματα στην αιχμή της διεθνούς επιστημονικής διερεύνησης.

ΠΕΡΙΣΣΟΤΕΡΕΣ ΠΑΗΡΟΦΟΡΙΕΣ:

Διαχείριση ρευμάτων κινούμενων αντικειμένων: <http://www.dblab.ece.ntua.gr/~kpatro/TSV/>

Συναφής διπλωματική εργασία (2011): <http://www.dblab.ece.ntua.gr/pubs/details.php?id=1625>