



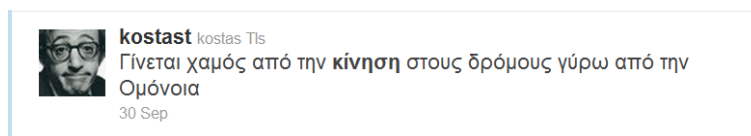
## Διπλωματική Εργασία 2011-2012

# Χρήση Χωροχρονικής Πληροφορίας από Δίκτυα Κοινωνικής Δικτύωσης για Εφαρμογές Διαχείρισης Κυκλοφορίας

### Εισαγωγή

Η εξάπλωση των ασύρματων δικτύων σε συνδυασμό με την εξέλιξη των κινητών τηλεφώνων σε υπερσύγχρονες συσκευές με ιδιαίτερες δυνατότητες έχει προάγει την ευρεία χρήση κοινωνικών δικτύων και υπηρεσιών από κινητούς χρήστες. Τέτοιοι χρήστες είναι μάλιστα περισσότερο ενεργοί, παράγοντας μεγαλύτερο όγκο δεδομένων και σε υψηλότερο ρυθμό από το μέσο χρήστη. Τα δεδομένα που παράγουν ποικίλουν από ακατέργαστα δεδομένα που λαμβάνονται από τις κινητές συσκευές (π.χ. συντεταγμένες GPS) μέχρι μηνύματα κειμένου, φωτογραφίες, βίντεο με τα οποία οι χρήστες ανταλλάσσουν νέα, εκφράζουν ιδέες και γνώμες, μοιράζονται τις θέσεις τους, τη διάθεση ή τα συναισθήματά τους.

Ένα από τα δημοφιλέστερα κοινωνικά δίκτυα είναι το twitter [6], που αυτοαποκαλείται υπηρεσία microblogging, επιτρέποντας στους χρήστες του να δημοσιοποιούν μηνύματα με μέγιστο μέγεθος 140 χαρακτήρων (tweets), αλλά και να ακολουθούν (follow) τα tweets των χρηστών με τους οποίους είναι συνδεδεμένοι. Εκεί οι κινητοί χρήστες συχνά δίνουν πληροφορίες σχετικές με την τοποθεσία στην οποία βρίσκονται, για εκδηλώσεις που λαμβάνουν χώρα στην πόλη, ακόμα και για την κίνηση στους δρόμους.



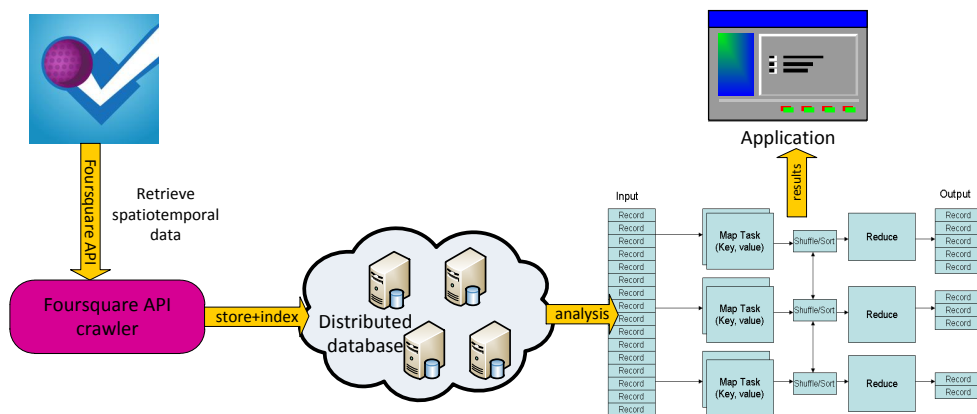
Σχήμα 1: Παράδειγμα tweet με πληροφορία κίνησης

Πέρα από τα κοινωνικά δίκτυα, ειδικά για κινητούς χρήστες υπάρχουν εξειδικευμένες υπηρεσίες, οι υπηρεσίες θέσης (Location Based Services or LBS). Με τον όρο αυτόν περιγράφεται το ευρύτερο σύνολο των εφαρμογών που προσπελάζονται μέσω κινητών συσκευών και κάνουν χρήση της γεωγραφικής τους θέσης. Μια από τις πιο ευρέως διαδεδομένες LBS είναι το Foursquare [4], που δίνει την δυνατότητα στους χρήστες να κάνουν check in στις διάφορες τοποθεσίες που επισκέπτονται (όπως εστιατόρια, καφετέ-

ριες, μουσεία κτλ), να μοιράζονται την πληροφορία αυτή με τους φίλους τους και να απολαμβάνουν προσωποποιημένες υπηρεσίες βάσει της θέσης τους.

Με πάνω από 200 εκατομμύρια χρήστες Twitter και 10 εκατομμύρια χρήστες Foursquare παγκοσμίως, είναι φανερό ότι υπάρχει ένας τεράστιος όγκος δεδομένων που παράγεται και που αποτελεί μια πολύτιμη πηγή γνώσης και πληροφορίας που μπορεί να συσχετιστεί χωρικά, χρονικά και κοινωνικά για την αναγνώριση συλλογικών καταστάσεων, π.χ. αυθόρμητων γεγονότων που μπορεί να συμβούν σε κάποιο σημείο και να επηρεάσουν την κίνηση.

Στην παρούσα διπλωματική θα ασχοληθούμε με τη συλλογή χωροχρονικών δεδομένων από κοινωνικά δίκτυα και LBS υπηρεσίες, την αποδοτική δεικτοδότηση και αποθήκευση τους και τη μαζική επεξεργασία τους για την εξαγωγή πληροφορίας για την κίνηση στους δρόμους, εξερευνώντας τεχνικές από την περιοχή των κατανεμημένων συστημάτων. Η πληροφορία αυτή είναι χρήσιμη τόσο σε ατομικό όσο και σε συλλογικό επίπεδο, όπως για παράδειγμα στη λήψη αποφάσεων από τοπικές αρχές ή οργανισμούς.



Σχήμα 2: Αρχιτεκτονική του συστήματος

## Σκοπός

Σκοπός της διπλωματικής είναι η ανάπτυξη ενός συστήματος συλλογής, αποθήκευσης και δεικτοδότησης χωροχρονικών και κοινωνικών δεδομένων που μοιράζονται οι κινητοί χρήστες μέσω Twitter και Foursquare αλλά και η επεξεργασία τους σε μεγάλη κλίμακα για την ανακάλυψη συλλογικών γεγονότων, την εξαγωγή συνθηκών κίνησης σε πραγματικό χρόνο και την έξυπνη διαχείριση κυκλοφοριακής ροής. Γι' αυτόν το σκοπό θα γίνει χρήση τεχνικών από την περιοχή των Κατανεμημένων Συστημάτων (Cloud/P2P, π.χ. [2, 3]). Η πορεία που θα ακολουθηθεί είναι η εξής:

1. Ανάπτυξη crawler για την συλλογή χρονικής και γεωγραφικής πληροφορίας (π.χ. γεωγραφικές συντεταγμένες) από tweets αλλά και από το Foursquare κάνοντας χρήση των αντίστοιχων APIs [1, 5], κατά προτίμηση σε Java.

2. Ανάπτυξη σε Java κατανεμημένου συστήματος αποθήκευσης των πληροφοριών αυτών με σκοπό την αποδοτική αναζήτησή τους.
3. Ανάλυση των πληροφοριών με χρήση κυκλοφοριακών μοντέλων κάνοντας χρήση τεχνικών κατανεμημένης επεξεργασίας (MapReduce framework).
4. Ανάπτυξη εφαρμογής για τον τελικό χρήστη, μέσω της οποίας θα μπορεί να βρίσκει πληροφορίες για την κίνηση στους δρόμους.
5. Πειραματική αξιολόγηση του συστήματος.

#### **Επικοινωνία:**

Νεκτάριος Κοζύρης, Αναπ. Καθηγητής nkoziris@cslab.ece.ntua.gr

Κατερίνα Δόκα, Μεταδ. Ερευνήτρια katerina@cslab.ece.ntua.gr

Δημήτριος Τσουμάκος, Επίκ. Καθηγητής dtsouma@cslab.ece.ntua.gr

#### **Βιβλιογραφία**

- [1] Foursquare API. <https://developer.foursquare.com/>. 2
- [2] The Apache HBase Project. <http://hbase.apache.org/>. 2
- [3] J. Dean and S. Ghemawat. Mapreduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1):107–113, 2008. 2
- [4] Foursquare. <https://foursquare.com/>. 1
- [5] Twitter API wiki. <http://apiwiki.twitter.com/>. 2
- [6] Twitter Webpage. <http://twitter.com/>. 1