



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΔΙΑ ΒΙΟΥ
ΜΑΘΗΣΗΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΕΡΕΥΝΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ «ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ»

ΕΘΝΙΚΟ ΣΤΡΑΤΗΓΙΚΟ ΠΛΑΙΣΙΟ ΑΝΑΦΟΡΑΣ
ΕΣΠΑ 2007-2013

ΔΡΑΣΗ «ΑΡΙΣΤΕΙΑ»

**Παραδοτέο 2.1 Απαιτήσεις συστήματος,
αρχιτεκτονική και σενάρια αξιολόγησης**

ΑΚΡΩΝΥΜΙΟ/ΚΩΔΙΚΟΣ:	INCEPTION/940
ΚΥΡΙΟΣ ΕΡΕΥΝΗΤΗΣ:	ΓΕΩΡΓΙΟΣ Δ. ΣΤΑΜΟΥΛΗΣ
ΦΟΡΕΑΣ ΥΠΟΔΟΧΗΣ:	ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ (ΟΠΑ)
ΣΥΝΤΑΚΤΕΣ ΠΑΡΑΔΟΤΕΟΥ:	Γ. Θάνος, Γ. Σταμούλης, Β. Καλογεράκη, Κ. Κουρκουμπέτης, Ε. Μαρκάκης, Ι. Μπούτσης (ΟΠΑ), Δ. Γουνόπουλος (ΕΚΠΑ)



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Ευρωπαϊκό Κοινωνικό Ταμείο

Περιεχόμενα

1. ΕΙΣΑΓΩΓΗ	3
2. ΒΑΣΙΚΕΣ ΑΠΑΙΤΗΣΕΙΣ - ΣΥΝΟΨΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ	5
3. ΠΑΡΟΥΣΙΑΣΗ ΤΗΣ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ	8
3.1 ΟΡΙΣΜΟΙ ΚΑΙ ΠΡΟΔΙΑΓΡΑΦΕΣ	8
3.1.1 <i>Χρήστες</i>	8
3.1.2 <i>Δεδομένα</i>	8
3.2 ΕΠΙΣΚΟΠΗΣΗ ΣΥΝΑΦΩΝ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΤΕΧΝΟΛΟΓΙΩΝ	9
3.2.1 <i>Επισκόπηση Συναφών Συστημάτων</i>	9
3.2.2 <i>MapReduce</i>	10
3.2.3 <i>MapReduce Frameworks</i>	10
3.2.4 <i>Publish/Subscribe</i>	11
3.3 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΣΥΣΤΗΜΑΤΟΣ	12
3.3.1 <i>Παραγωγή και αποθήκευση Δεδομένων</i>	12
3.4 ΑΝΑΚΤΗΣΗ ΔΕΔΟΜΕΝΩΝ	13
3.5 ΕΝΔΕΙΚΤΙΚΟ ΣΕΝΑΡΙΟ ΧΡΗΣΗΣ	14
3.5.1 <i>Ενσωμάτωση Μηχανισμών Αγοράς</i>	15
4. ΠΑΡΟΥΣΙΑΣΗ ΤΗΣ ΥΠΗΡΕΣΙΑΣ ΑΓΟΡΑΣ (MARKET SERVICE) ΚΑΙ ΤΩΝ ΜΗΧΑΝΙΣΜΩΝ ΚΙΝΗΤΡΩΝ	16
4.1 ΠΙΘΑΝΑ ΣΕΝΑΡΙΑ ΕΠΙΔΕΙΞΗΣ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ	17
4.2 ΕΠΙΛΟΓΗ ΚΑΙ ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΒΑΣΙΚΟΥ ΣΕΝΑΡΙΟΥ	18
4.3 ΠΑΡΑΔΕΙΓΜΑ ΥΛΟΠΟΙΗΣΗΣ ΜΗΧΑΝΙΣΜΟΥ ΤΗΣ ΥΠΗΡΕΣΙΑΣ ΑΓΟΡΑΣ ΓΙΑ ΤΟ ΒΑΣΙΚΟ ΣΕΝΑΡΙΟ	20
5. ΣΕΝΑΡΙΑ ΑΞΙΟΛΟΓΗΣΗΣ ΚΑΙ ΚΡΙΤΗΡΙΑ ΕΠΙΤΥΧΙΑΣ	23
6. ΣΥΝΟΨΗ - ΣΥΜΠΕΡΑΣΜΑΤΑ	24
7. ΑΝΑΦΟΡΕΣ	26

1. Εισαγωγή

Η νέα γενιά των προσωπικών συσκευών κινητής τηλεφωνίας έχει αλλάξει σημαντικά τον τρόπο που αλληλεπιδρούν οι άνθρωποι μεταξύ τους, με τη χρήση ισχυρών, προγραμματιζόμενων συσκευών που είναι εξοπλισμένες με μια σειρά από εξελιγμένες δυνατότητες ανίχνευσης και αίσθησης και με αντίστοιχο εξοπλισμό, όπως κάμερες, αισθητήρες GPS, και επιταχυνσιόμετρα. Οι εξελίξεις στην δικτύωση αισθητήρων και η διαθεσιμότητα των συσκευών με αισθητήρες χαμηλού κόστους έχουν οδηγήσει στην ενσωμάτωση των αισθητήρων για την παρακολούθηση του φυσικού κόσμου και την εξαγωγή χρήσιμων δεδομένων και συμπερασμάτων από αυτά, τα οποία στη συνέχεια μπορούν να αξιοποιηθούν κατάλληλα. Αυτό ήδη είναι εμφανές σε ένα ευρύ φάσμα ζωτικής σημασίας τομείς όπως οι μεταφορές, περιβαλλοντικές εφαρμογές κτλ. Έτσι, η δραματική αύξηση της λήψης δεδομένων και της ανταλλαγής τους έχει δημιουργήσει μια επιτακτική ανάγκη για συντονισμό των εμπλεκόμενων μέσω της ανάπτυξης νέων μηχανισμών για την υποστήριξη μιας κοινότητας/περιβάλλοντος αισθητήρων, όπου πολλοί χρήστες, επιχειρήσεις και εφαρμογές θα αλληλεπιδρούν μεταξύ τους και θα ανταλλάσσουν δεδομένα λαμβάνοντας υπόψη και οικονομικά δεδομένα και ζητήματα, όπως κόστος, ανταποδοτικό όφελος κτλ.

Οι τρέχουσες ερευνητικές πρωτοβουλίες έχουν επικεντρωθεί στην επίλυση των τεχνικών προκλήσεων της δικτύωσης και της τεχνολογίας πληροφοριών, όπως η δυνατότητα δημιουργίας χαμηλής ισχύος, και χαμηλού εύρους ζώνης ασυρμάτων δικτύων από αισθητήρες, η απαιτούμενη ενέργεια, η διαχείριση των πόρων (π.χ. των κινητών συσκευών) κλπ., ορισμένα από τα οποία προκύπτουν και από φυσικούς περιορισμούς. Ωστόσο, το θεμελιώδες οικονομικό ζήτημα, δηλ. γιατί οι χρήστες να επιθυμούν να μοιράζονται και να ανταλλάσσουν τις πληροφορίες αυτές, μια διαδικασία που μπορεί να είναι δαπανηρή γι' αυτούς, αλλά ταυτόχρονα και επωφελής για άλλους (ή και για τον ίδιο) δεν έχει αντιμετωπιστεί. Για τη μελέτη αυτή είναι απαραίτητο να επιλεγεί και η απαραίτητη τεχνολογία που να δύναται να υποστηρίξει τις ανταλλαγές αυτές μέσα σε ένα βιώσιμο σύστημα που με ενδεχόμενη μεγάλη αύξηση της κλίμακας και του πληθυσμού των χρηστών.

Αντικείμενο του παρόντος έργου INCEPTION είναι να αξιοποιήσει αυτό το φαινόμενο ομαδικής συλλογής δεδομένων και τις παράλληλες τεχνολογικές εξελίξεις, και να δημιουργήσει ένα νέο πλαίσιο και τεχνολογικό πρωτότυπο για το πώς οι χρήστες αλληλεπιδρούν μεταξύ τους σε έναν εικονικό χώρο ενστερνιζόμενοι τις αρχές και τα εργαλεία της οικονομικής επιστήμης και τις τεχνολογικές δυνατότητες. Ο τελικός στόχος του έργου INCEPTION είναι ο σχεδιασμός και η ανάπτυξη **μιας ανοικτής πλατφόρμας λογισμικού βασιζόμενης στην παροχή οικονομικών κινήτρων** για την συνεργασία μεταξύ αισθητήρων, κινητών και «έξυπνων» συσκευών και χρηστών-παρόχων δεδομένων και εφαρμογών.

Το παρόν παραδοτέο αποτελεί το πρώτο σημαντικό ορόσημο στην πορεία των εργασιών του έργου προς τον τελικό στόχο του, γιατί παρουσιάζει τις απαιτήσεις που καλείται να υποστηρίξει η υπό ανάπτυξη πλατφόρμα (σε οικονομικό και τεχνολογικό επίπεδο), την αρχιτεκτονική της πλατφόρμας αυτής, περιλαμβανομένης και της λογικής στην οποία θα βασίζεται η οικονομική βελτιστοποίηση που θα συντελείται στα πλαίσια της αρχιτεκτονικής, και τα σενάρια για την αξιολόγησης της.

Η δομή του παρόντος παραδοτέου έχει ως εξής: Το 2^ο κεφάλαιο παρουσιάζει μία σύνοψη της αρχιτεκτονικής μαζί με τις βασικές απαιτήσεις υλοποίησης της, όπως και μία περιγραφή της αλληλεπίδρασης των επιμέρους υποσυστημάτων. Η αρχιτεκτονική παρουσιάζεται λεπτομερώς στο επόμενο Κεφάλαιο (3^ο) όπως και ένα ενδεικτικό σενάριο χρήσης της. Ένα από τα βασικά μέρη της αρχιτεκτονικής είναι η υπηρεσία αγοράς (market service) και των μηχανισμών κινήτρων. Αυτά παρουσιάζονται και αναλύονται λεπτομερώς στο 4^ο Κεφάλαιο, όπου γίνεται και μια πλήρης ταξινόμηση των πιθανών περιπτώσεων ζήτησης για πληροφορία και χαρακτηριστικών συνέργειας των μηχανισμών κινήτρων. Επιπρόσθετα, περιγράφονται τα πιθανά σενάρια επίδειξης του συστήματος, γίνεται η επιλογή ενός από αυτά ως βασικού (με βάση το ερευνητικό ενδιαφέρον που παρουσιάζει αλλά και την αντιπροσωπευτικότητα των χαρακτηριστικών του) και παρουσιάζεται ένα παράδειγμα υλοποίησης μηχανισμού της υπηρεσίας αγοράς για το συγκεκριμένο σενάριο. Το 5^ο Κεφάλαιο περιγράφει τα βασικά σενάρια αξιολόγησης και τα κριτήρια επιτυχίας του συστήματος, ενώ τέλος το 6^ο Κεφάλαιο συνοψίζει τα βασικά συμπεράσματα του παραδοτέου.

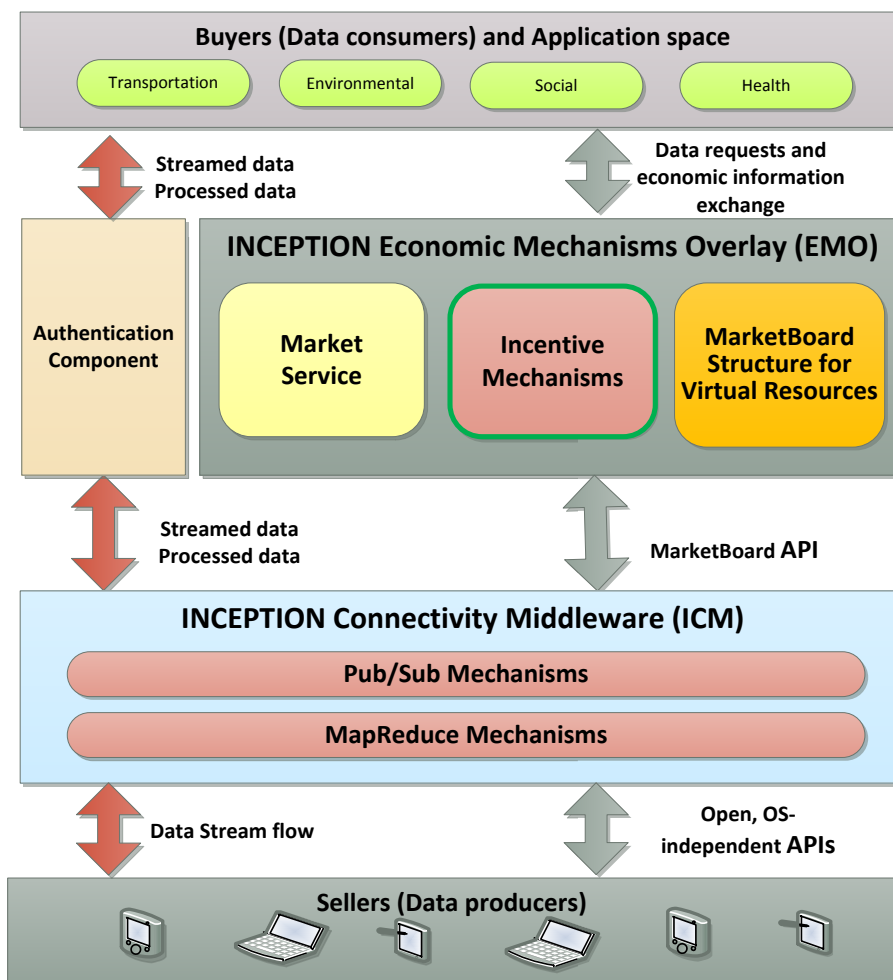
2. Βασικές Απαιτήσεις - Σύνοψη Αρχιτεκτονικής

Ο στόχος του INCEPTION είναι να διερευνήσει τα κοινωνικοοικονομικά κίνητρα ενός περιβάλλοντος χρηστών που συμμετέχουν σε δίκτυα αισθητήρων, και αλληλεπιδρούν μεταξύ τους και με άλλες κοινωνικές ομάδες με σκοπό να επιτύχουν ανάπτυξη και οικονομική βιωσιμότητα. Στην προσέγγισή μας, οι χρήστες που διαθέτουν αισθητήρες αντιμετωπίζονται ως οικονομικές οντότητες: υφίστανται κόστος όταν συλλέγουν και μεταδίδουν δεδομένα, αλλά ωφελούνται από τη χρήση εφαρμογών που βασίζονται στην συλλογή πληροφοριών που παρέχονται σε συνδυασμό και από τα άλλα μέλη της αντίστοιχης κοινότητας.

Στόχος μας είναι να αναπτύξουμε ένα νέο σύστημα που να προωθεί τη συμβατότητα των κινήτρων για την ανταλλαγή πληροφοριών μεταξύ των χρηστών και των εφαρμογών. Οι βασικές απαιτήσεις του παρόντος συστήματος είναι οι εξής

1. Οι εφαρμογές που αξιοποιούν τα δεδομένα που δύνανται να συνεισφέρουν οι χρήστες πρέπει να βασίζονται στα κίνητρα των χρηστών, και συνεπώς, να παρέχουν τους τις κατάλληλες ανταμοιβές, όταν η συμπεριφορά τους στη συγκέντρωση δεδομένων είναι χρήσιμη στους άλλους, αυξάνοντας έτσι την οικονομική αποδοτικότητα και τη βελτιστοποίηση των συνολικά διαθέσιμων δεδομένων στο σύστημα, και συντελώντας στην οικονομική βιωσιμότητα του. Οι χρήστες θα πρέπει να ενθαρρύνονται να παρέχουν τα δεδομένα εκείνα που είναι πιο χρήσιμα για τις εφαρμογές και τις κοινωνικές ομάδες που είναι ενεργές, αποφεύγοντας παράλληλα την παραγωγή αυτών που είναι δαπανηρά και λιγότερο χρήσιμα. Αυτή η δυναμική πτυχή της προστιθέμενης αξίας των πληροφοριών είναι ζωτικής σημασίας και μπορεί να αντιμετωπιστεί μόνο μέσα σε ένα οικονομικό πλαίσιο, όπως αυτό που προτείνει το έργο INCEPTION.
2. Σε τεχνολογικό επίπεδο πρέπει το σύστημα να επιτρέπει την αλληλεπίδραση των χρηστών με κινητές συσκευές που χρησιμοποιούν λειτουργικά συστήματα ευρείας χρήσης και αποδοχής όπως το Android. Επίσης, το ενδιάμεσο λογισμικό (βλέπε κατωτέρω) θα πρέπει να παρέχει πρόσβαση σε ένα βασικό αριθμό αισθητήρων που υποστηρίζουν οι σύγχρονες έξυπνες συσκευές, όπως για παράδειγμα πρόσβαση στον αισθητήρα GPS. Επιπρόσθετα, το σύστημα θα πρέπει να εξασφαλίζει την ασφαλή και εγγυημένη ανώνυμη επικοινωνία με τον χρήστη, ώστε να προστατεύει τα προσωπικά του στοιχεία αλλά και δεδομένα από τρίτους και έτσι να αίρονται τυχόν ενδοιασμοί του χρήστη για την χρήση του συστήματος INCEPTION.
3. Σε κοινωνικό επίπεδο, η κύρια απαίτηση είναι η προαναφερθείσα για προστασία της ιδιωτικότητας (privacy) του χρήστη, η οποία αποτελεί ένα σημαντικό ζήτημα γενικώς για τα νέα συστήματα επεξεργασίας δεδομένων, και θέτει αντίστοιχες τεχνολογικές απαιτήσεις στην υλοποίησή τους.

Για να επιτευχθούν τα παραπάνω η ερευνητική ομάδα του έργου προτείνει την υλοποίηση μας αρχιτεκτονικής όπως παρουσιάζεται στο Σχήμα 1. Σε αυτή παρουσιάζονται επίσης οι διεπαφές του συστήματος και οι αλληλεπιδράσεις μεταξύ των υποσυστημάτων, τα οποία είναι τα εξής:



Σχήμα 1. Η αρχιτεκτονική του INCEPTION

- Υποσύστημα Οικονομικών Μηχανισμών (Economic Mechanisms Overlay), που περιλαμβάνει και εκτελεί τους μηχανισμούς αγοράς (Κεφάλαιο 4) και διαχειρίζεται τις περιγραφές των διατιθεμένων από χρήστες δεδομένων τα οποία εκχωρούνται μέσω των μηχανισμών αυτών (Κεφάλαιο 3).
- Ενδιάμεσο Λογισμικό (Connectivity Middleware), το οποίο διαχειρίζεται τα ίδια τα διατιθέμενα από χρήστες δεδομένα (Κεφάλαιο 3).

Σημειώτέον ότι σε μια υλοποίηση της αρχιτεκτονικής για πρακτική εφαρμογή, τα ανωτέρω υποσυστήματα θα συμπληρώνονταν από το Λογισμικό Αυθεντικοποίησης (authentication component) για πιστοποίηση της αυθεντικότητας του χρήστη (ταυτότητα, δικαιώματα χρήσης κτλ.) σε κάθε συναλλαγή του με το σύστημα.

Συνοπτικά, ένας χρήστης του INCEPTION αλληλεπιδρά σε τρία επίπεδα με την αρχιτεκτονική:

1. Εγγραφή: Ο χρήστης-πάροχος δεδομένων συνδέεται σε μια διαδικτυακή πύλη (portal) χρησιμοποιώντας Open APIs (Application Programming Interfaces) μέσω μιας έξυπνης συσκευής. Κατά τη διαδικασία εγγραφής, ο χρήστης εγκαθιστά το κατάλληλο πρόγραμμα που

θα επιτρέψει τη χρήση του συστήματος.

2. Αλληλεπίδραση με την Υπηρεσία Αγοράς: Ο χρήστης καταχωρεί και περιγράφει τη διαθεσιμότητα των δεδομένων του μέσω του συγκεκριμένου API. Οι περιγραφές στη συνέχεια φορτώνονται στην δομή της Αγοράς (Marketboard), και είναι διαθέσιμες για όλους τους χρήστες του συστήματος.

Ο χρήστης μπορεί ο ίδιος να ορίσει μία τιμή για τα δεδομένα, ή να του προτείνει το σύστημα μία με βάση τις άλλες προσφορές. Ένας χρήστης-αγοραστής δεδομένων από την άλλη χρησιμοποιεί το API ως μηχανισμό αναζήτησης για να μάθει τι διαθέσιμα δεδομένα υπάρχουν που ταιριάζουν με τα κριτήρια που έθεσε και ενδεχομένως να τα ζητήσει. Η υπηρεσία αγοράς αναλαμβάνει να ικανοποιήσει τις αιτήσεις των αγοραστών και να τους χρεώσει ανάλογα, κατά τον βέλτιστο τρόπο λαμβάνοντας υπ' όψιν τις μεθοδολογίες που θα οριστούν στην Ενότητα Εργασίας 3.

3. Η ανταλλαγή δεδομένων: Όταν μια συναλλαγή αποφασίζεται από την υπηρεσία αγοράς, η πραγματική ανταλλαγή δεδομένων πραγματοποιείται μέσω ενός ενδιάμεσου λογισμικού που υλοποιεί τις συναλλαγές που αποφασίστηκαν.

Τα παραπάνω στάδια θα αναλυθούν σε μεγαλύτερη λεπτομέρεια στις επιμέρους ενότητες κατά την περιγραφή των υποσυστημάτων της προτεινόμενης αρχιτεκτονικής του έργου.

3. Παρουσίαση της Αρχιτεκτονικής

Στο παρόν κεφάλαιο γίνεται η παρουσίαση της αρχιτεκτονικής της συστήματος, αποτελούμενης από τη Δομή MarketBoard και το ενδιάμεσο λειτουργικό (βλέπε Κεφάλαιο 2), σύμφωνα και με ένα απλό σενάριο επίδειξης του συστήματος το οποίο και αναπτύσσεται στο τέλος του παρόντος κεφαλαίου. Σημειωτέον ότι τα κύρια σενάρια επίδειξης του συστήματος αναπτύσσονται στο Κεφάλαιο 4.

3.1 Ορισμοί και Προδιαγραφές

3.1.1 Χρήστες

Έστω ότι έχουμε ένα σύστημα που αποτελείται από N χρήστες και οι χρήστες αυτοί συμμετέχουν μέσω των συσκευών τους (κινητά, tablets κτλ). Για την αποδοτική λειτουργία του συστήματος απαιτείται ένας ικανοποιητικός αριθμός χρηστών και όσο αυτός αυξάνεται τόσο θα αυξάνεται και η ακρίβεια των αποτελεσμάτων του συστήματος, λόγω του διαθέσιμου συνόλου των δεδομένων για κάθε γεωγραφική περιοχή και χρονική στιγμή.

Τα δεδομένα των χρηστών παράγονται από τους αισθητήρες των συσκευών τους και αποθηκεύονται τοπικά σε κάθε συσκευή, ενώ το σύστημα διατηρεί μόνο αναφορές για κάθε πλειάδα δεδομένων. Τις αναφορές αυτές μπορεί να χρησιμοποιήσουν οι χρήστες του MarketBoard για να επιλέξουν τα δεδομένα που θέλουν. Εφόσον οι χρήστες δεν είναι συνδεδεμένοι συνεχώς στο σύστημα λόγω της διακοπτόμενης συνδεσιμότητας των φορητών συσκευών αντίστοιχα ούτε τα δεδομένα τους θα προσφέρονται συνεχώς στο MarketBoard.

3.1.2 Δεδομένα

Κάθε συσκευή (κινητό, tablet, κτλ) μπορεί να παράγει πλειάδες δεδομένων μέσω των αισθητήρων της (πχ αισθητήρες GPS, επιταχυνσιόμετρο, κάμερα, μικρόφωνο κτλ) οι οποίες ονομάζονται Application Data Units (ADU).

Κάθε πλειάδα δεδομένων που παράγεται από τον χρήστη i για μία χρονική στιγμή t ορίζεται ως $ADU_i(t)$ και οι πλειάδες αυτές αποθηκεύονται τοπικά στη συσκευή του χρήστη που τις παράγει.

Η μορφή και το σύνολο των δεδομένων εξαρτάται από την εφαρμογή με τη γενική μορφή να είναι $\langle ID, id_χρήστη, latitude, longitude, χρονοσφραγίδα, δεδομένα εφαρμογής \rangle$ όπου:

- Το ID αναφέρεται στο μοναδικό κωδικό του ADU που δημιουργείται από τον κατακερματισμό της θέσης και της χρονοσφραγίδας (θεωρούμε ότι η ακρίβεια της θέσης είναι τέτοια ώστε να μην μπορούν να υπάρχουν 2 χρήστες στο ίδιο σημείο ταυτόχρονα)
- Το $id_χρήστη$ είναι το id του χρήστη στο σύστημα
- τα $latitude, longitude$ προσδιορίζουν τη γεωγραφική θέση της πλειάδας
- η χρονοσφραγίδα αναφέρεται στο χρόνο παραγωγής των δεδομένων
- τα δεδομένα εφαρμογής αναφέρονται σε συγκεκριμένα δεδομένα της εφαρμογής που μας ενδιαφέρει

Παραδείγματα ADU είναι:

- <ID, id_χρήστη, latitude, longitude, χρονοσφραγίδα, δεδομένα video> για μια εφαρμογή παρακολούθησης κίνησης μέσω κάμερας,
- < ID, id_χρήστη, latitude, longitude, χρονοσφραγίδα, δεδομένα επιταχυνσιόμετρου, δείγματα μικροφώνου> για μία εφαρμογή παρακολούθησης για σεισμούς.

Το μέγεθος των ADU διαφέρει ανάλογα με την εφαρμογή και κάθε ADU μπορεί να συνδυάζει πολλούς τύπους δεδομένων με διαφορετικά χαρακτηριστικά, όπως γίνεται εμφανές και από τα παραδείγματα.

Επερωτήσεις

Κάθε χρήστης μπορεί να κάνει επερωτήσεις (queries) στο σύστημα για να πάρει τα δεδομένα που έχουν παραχθεί από τους χρήστες (ADU). Έτσι, κάθε επερώτηση θα πρέπει να περιέχει ένα προσδιορισμό του συνόλου των πλειάδων που θα επιστραφούν (πχ. Γεωγραφικό προσδιορισμό – «Δώσε μου τις πλειάδες που αναφέρονται σε μια συγκεκριμένη γεωγραφική περιοχή» καθώς και περιορισμούς όπως το μέγιστο αριθμό των πλειάδων που θα επιστραφούν σαν αποτέλεσμα.

Τα δεδομένα που επιστρέφονται αναφέρονται:

- είτε σε ένα υποσύνολο των πλειάδων κάθε χρήστη (ιστορικά δεδομένα) πχ. «επέστρεψε το σύνολο των πλειάδων που έχεις σε αυτή τη γεωγραφική περιοχή»
- είτε στην πιο πρόσφατη πλειάδα κάθε χρήστη «realtime δεδομένα»

Σε κάθε επερώτηση δίνεται η δυνατότητα στο χρήστη να ζητάει ένα υποσύνολο κάθε πλειάδας. Για παράδειγμα, για τις πλειάδες που περιγράψαμε παραπάνω για την εφαρμογή παρακολούθησης για σεισμούς μπορούμε να ζητάμε μόνο τις μετρήσεις του επιταχυνσιόμετρου σε μια δεδομένη χρονική στιγμή.

3.2 Επισκόπηση Συναφών Αρχιτεκτονικών Τεχνολογιών

Στο παρόν υποκεφάλαιο γίνεται η επισκόπηση των κυριότερων τεχνολογιών που σχετίζονται με την υπό ανάπτυξη αρχιτεκτονική

3.2.1 Επισκόπηση Συναφών Συστημάτων

Στη βιβλιογραφία μπορούν να βρεθούν διάφορα συστήματα τα οποία χρησιμοποιούν κινητά για να παράγουν δεδομένα και να τα χρησιμοποιήσουν σε διάφορους τομείς. Σε αυτά τα συστήματα οι χρήστες συνήθως παράγουν δεδομένα μέσω των αισθητήρων των κινητών συσκευών και κάνουν κάποια μικρή επεξεργασία πάνω από αυτά μέσω της συσκευής τους. Τα δεδομένα στη συνέχεια μεταφέρονται σε κάποιο άλλο κόμβο για να συγχωνευθούν και να παραχθεί το επιθυμητό αποτέλεσμα. Παραδείγματα τέτοιων συστημάτων είναι τα Vtrack [1] και MetroSense [2] στον τομέα των μεταφορών που έχουν ως σκοπό την ανίχνευση κίνησης και την εκτίμηση του χρόνου μίας διαδρομής σε πραγματικό χρόνο, χρησιμοποιώντας τα δεδομένα των χρηστών που βρίσκονται στην εκάστοτε περιοχή. Αντίστοιχα το GreenGPS [3] χρησιμοποιεί τα δεδομένα των χρηστών για να εξάγει τις πιο αποδοτικές διαδρομές από άποψη κατανάλωσης καυσίμων. Αντίστοιχα συστήματα υπάρχουν για διάφορους τομείς όπως για τον προσδιορισμό περιοχών καλής WiFi συνδεσιμότητας [4] καθώς και για

συστήματα περιβαλλοντικής παρακολούθησης [5] και ανίχνευσης σεισμών [6].

3.2.2 MapReduce

Το MapReduce είναι ένα μοντέλο που επιτρέπει την παράλληλη επεξεργασία μεγάλων όγκων δεδομένων. Η αρχιτεκτονική MapReduce στηρίζεται στη χρήση δύο συναρτήσεων:

- `map(key,value) -> [(key2, value2)]`
- `reduce(key2,[value2]) -> [final_value]`

Η «Map» συνάρτηση επεξεργάζεται ένα ζεύγος key/value και παράγει ένα ενδιάμεσο ζεύγος key/value. Η είσοδος στη συνάρτηση map έχει τη μορφή (κλειδί, τιμή) και μετατρέπει κάθε τέτοιο ζευγάρι σε ένα άλλο ζευγάρι (κλειδί2, τιμή2) ως έξοδο της συνάρτησης. Η Map μπορεί να εκτελείται παράλληλα, πάνω σε διαφορετική είσοδο δεδομένων, σε διαφορετικούς κόμβους.

Η «Reduce» συνάρτηση συγχωνεύει όλα τα ενδιάμεσα values που έχουν παραχθεί από τη Map και σχετίζονται με το ίδιο κλειδί και παράγει τα τελικά αποτελέσματα. Έτσι, υπολογίζει μια τελική τιμή για το κλειδί, επεξεργάζοντας τη λίστα των τιμών που αντιστοιχούν σε αυτό το κλειδί. Για να επιτευχθεί αυτό, η επεξεργασία της συνάρτησης reduce γίνεται αφού έχει τελειώσει η επεξεργασία όλων των map συναρτήσεων.

Το πλεονέκτημα της αρχιτεκτονικής MapReduce είναι ότι επιτρέπει την κατανεμημένη επεξεργασία των εργασιών μέσω των συναρτήσεων «map» και «reduce». Υπό την προϋπόθεση ότι κάθε «map» συνάρτηση είναι ανεξάρτητη από την άλλη, όλες οι «map» συναρτήσεις μπορούν να εκτελούνται παράλληλα ώστε να ελαχιστοποιηθεί ο χρόνος εκτέλεσης.

Το MapReduce είναι μία αρχιτεκτονική που έχει διαδοθεί πολύ πρόσφατα λόγω της ευελιξίας που προσφέρει και τη δυνατότητα να μειώσει δραστικά το χρόνο εκτέλεσης, σε χρονοβόρες διαδικασίες οι οποίες μπορούν να εκτελούνται παράλληλα. Έτσι έχει χρησιμοποιηθεί σε πολλά συστήματα για να βελτιώσει την απόδοσή τους όπως: στον υπολογισμό Theta-Joins [7][8], στο να δέχεται εργασίες από τους χρήστες ή συντάκτες γλωσσών υψηλού επιπέδου και να τις εκτελεί παράλληλα σε ένα cluster [9], για την εύρεση του πυκνότερου υπογράφου σε ένα μοντέλο ροής δεδομένων [10], για την υλοποίηση της κλειστότητας ενός γραφήματος RDF [11], για την κλιμάκωση γενετικών αλγορίθμων [12] κτλ.

3.2.3 MapReduce Frameworks

Στην αρχιτεκτονική μας θα χρησιμοποιήσουμε δύο πλαίσια υλοποίησης frameworks της αρχιτεκτονικής Map Reduce, το Hadoop¹ και το MISCO [14].

Το Hadoop της Apache είναι ένα framework ανοικτού κώδικα που βασίζεται στις δημοσιεύσεις της Google για το MapReduce και το Google File System (GFS). Το Hadoop επιτρέπει την κατανεμημένη επεξεργασία μεγάλων συνόλων δεδομένων σε clusters υπολογιστών με τη χρήση απλών μοντέλων προγραμματισμού. Χρησιμοποιεί το MapReduce ώστε να χωρίσει κάθε εργασία σε πολλά μικρά κομμάτια, καθένα από τα οποία μπορεί να

¹ <http://hadoop.apache.org/>

εκτελεστεί ή να ξαναεκτελεστεί σε κάθε κόμβο του συστήματος. Επιπλέον, παρέχει ένα καταμεμημένο σύστημα αρχείων που αποθηκεύει δεδομένα σχετικά με τους κόμβους υπολογιστικής, παρέχοντας πολύ υψηλό συνολικό εύρος ζώνης σε όλη την συστάδα.

Το Hadoop έχει σχεδιαστεί για να δουλεύει αποδοτικά από έναν υπολογιστή μέχρι χιλιάδες υπολογιστές. Μπορεί να εντοπίζει και να χειρίζεται αποτυχίες σε επίπεδο εφαρμογών, παρέχοντας έτσι εξαιρετική διαθεσιμότητα σε clusters υπολογιστών, όπου κάθε υπολογιστής μπορεί να είναι επιρρεπής σε αποτυχίες.

Το Hadoop είναι γραμμένο στη γλώσσα προγραμματισμού Java. Επίσης υποστηρίζει τη χρήση οποιασδήποτε γλώσσας προγραμματισμού μέσω «streaming» για να υλοποιηθούν οι συναρτήσεις map και reduce.

Το MISC0 είναι ένα framework σχεδιασμένο ώστε να υλοποιεί την αρχιτεκτονική Map Reduce σε κινητές συσκευές. Έτσι η ιδιαιτερότητά του είναι πως έχει σχεδιαστεί ώστε να δουλεύει αποδοτικά σε κινητές συσκευές οι οποίες έχουν περιορισμένους πόρους (cpu, ram, κτλ). Το MISC0 παρέχει μια ποικιλία πολιτικών δρομολόγησης για την επιλογή των εργατών που παίρνουν υπ' όψιν παραμέτρους όπως τη διαθεσιμότητα των συσκευών, το χρονικό όριο εκτέλεσης μίας εργασίας κτλ. Επίσης το MISC0 υποστηρίζει την ασθενή συνδεσιμότητα η οποία είναι ιδανική για δίκτυα κινητών συσκευών.

Εκτός από το MISC0 υπάρχουν και άλλα συστήματα στη βιβλιογραφία που προσπαθούν να υλοποιήσουν το Map Reduce πάνω από κινητές συσκευές με πιο χαρακτηριστικό το Hyrax [13]. Το Hyrax είναι μια πλατφόρμα που προέρχεται από το Hadoop ώστε να υποστηρίξει το cloud computing σε συσκευές Android. Η επιλογή μας όμως να χρησιμοποιήσουμε το MISC0 βασίζεται στο γεγονός ότι το MISC0 έχει δημιουργηθεί και βελτιστοποιηθεί εξ αρχής ώστε να δουλεύει αποδοτικά πάνω σε κινητές συσκευές αντί να έχει προσαρμοστεί σε αυτές.

3.2.4 Publish/Subscribe

Η αρχιτεκτονική Publish-Subscribe (Pub / sub) (βλέπε για παράδειγμα [15]) είναι ένα πρότυπο ανταλλαγής μηνυμάτων όπου οι αποστολείς των μηνυμάτων, που ονομάζονται εκδότες (publishers) δεν στέλνουν τα μηνύματα τους απευθείας στους παραλήπτες, οι οποίοι ονομάζονται συνδρομητές (subscribers). Τα μηνύματα που δημοσιεύονται από τους εκδότες κατηγοριοποιούνται σε τάξεις, χωρίς να γνωρίζουν ενδεχομένως τους συνδρομητές που μπορεί να υπάρχουν σε κάθε τάξη. Αντίστοιχα, οι συνδρομητές εκφράζουν ενδιαφέρον για μία ή περισσότερες τάξεις, και λαμβάνουν μόνο τα μηνύματα των τάξεων για τις οποίες έχουν εκφράσει ενδιαφέρον.

Φιλτράρισμα μηνυμάτων

Στην αρχιτεκτονική pub / sub, οι συνδρομητές λαμβάνουν συνήθως μόνο ένα υποσύνολο του συνόλου των δημοσιευμένων μηνυμάτων. Η διαδικασία της επιλογής των μηνυμάτων για την υποδοχή και την επεξεργασία καλείται «φιλτράρισμα» (filtering). Υπάρχουν δύο κοινές μορφές φιλτραρίσματος: (i) φιλτράρισμα με βάση το θέμα (topic) και (ii) με βάση το περιεχόμενο.

Σε ένα σύστημα όπου εκτελείται φιλτράρισμα με βάση το θέμα, τα μηνύματα δημοσιεύονται σε "θέματα", που ονομάζονται λογικά κανάλια. Οι συνδρομητές σε τέτοια συστήματα θα λάβουν όλα τα μηνύματα που δημοσιεύτηκαν στα θέματα στα οποία έχουν εγγραφεί. Ο εκδότης είναι υπεύθυνος για τον καθορισμό των θεμάτων στα οποία μπορούν να εγγραφούν

οι συνδρομητές.

Σε ένα σύστημα όπου εκτελείται φιλτράρισμα με βάση το περιεχόμενο, κάθε μήνυμα παραδίδεται σε έναν συνδρομητή, εάν τα χαρακτηριστικά ή το περιεχόμενο του μηνύματος αντιστοιχεί στους περιορισμούς που ορίζονται από τον συνδρομητή.

Μερικά συστήματα υποστηρίζουν ένα συνδυασμό των δύο μορφών φιλτραρίσματος. Οι εκδότες δημοσιεύουν τα μηνύματα σε ένα θέμα ενώ οι συνδρομητές εγγράφονται σε ένα ή περισσότερα θέματα με βάση το περιεχόμενο.

Πλεονεκτήματα

Ένα πλεονέκτημα της αρχιτεκτονικής pub/sub είναι πως οι εκδότες δεν χρειάζεται να γνωρίζουν την ύπαρξη των συνδρομητών του, διατηρώντας έτσι «χαλαρή» σύνδεση μαζί τους. Ένα δεύτερο πλεονέκτημα είναι πως το μοντέλο αυτό παρέχει την ευκαιρία για καλύτερη επεκτασιμότητα από τις παραδοσιακές client-server αρχιτεκτονικές, λόγω διαφόρων χαρακτηριστικών όπως π.χ. της παράλληλης λειτουργίας πολλαπλών εκδοτών.

3.3 Αρχιτεκτονική Συστήματος

Στο σύστημα μας κάθε χρήστης μπορεί να έχει δύο διακριτούς ρόλους:

- Να προσφέρει δεδομένα
- Να κάνει επερωτήσεις για δεδομένα

Φυσικά ένας χρήστης μπορεί να χρησιμοποιεί και τους δύο ρόλους ώστε να παράγει αλλά και να ζητάει δεδομένα από το σύστημα.

Χρησιμοποιούμε δύο διαφορετικούς μηχανισμούς για αυτούς τους δύο ρόλους ώστε να εκτελούνται οι απαραίτητες διαδικασίες αποδοτικά.

Ουσιαστικά η παραγωγή των δεδομένων και η δημοσίευση των αναφορών στο MarketBoard αναφέρεται στο Publish ενώ η ανάκτησή τους όταν υπάρχει ζήτηση από κάποιους χρήστες αναφέρεται στο Subscribe μίας Pub/Sub αρχιτεκτονικής. Έτσι στηριζόμαστε στο Map Reduce τόσο για την αποθήκευση όσο και για την ανάκτηση των δεδομένων, και στο Pub/Sub σαν γενική αρχιτεκτονική.

3.3.1 Παραγωγή και αποθήκευση Δεδομένων

Κάθε συσκευή παράγει και αποθηκεύει δεδομένα τοπικά, μέσω των αισθητήρων της, και το σύστημα κρατάει μόνο αναφορές αυτών των δεδομένων ώστε να μπορούν να ανακτηθούν μέσω των επερωτήσεων.

Για να προσφέρονται τα δεδομένα από τους χρήστες στο σύστημα μας, χρησιμοποιούμε το MISCO το οποίο είναι ένα framework για MapReduce, εξειδικευμένο σε κινητές συσκευές.

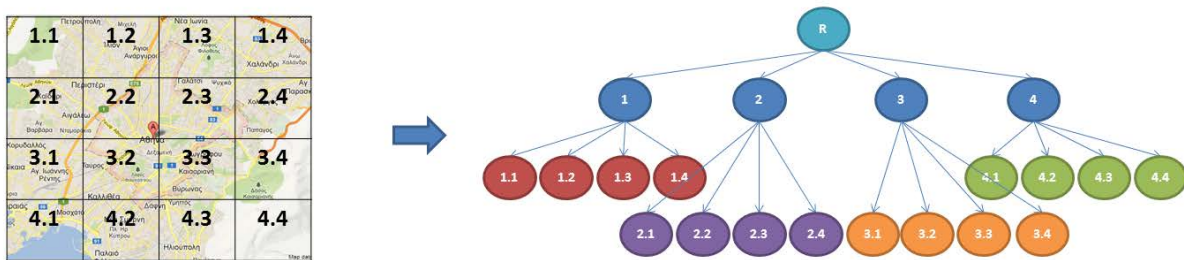
Η συνάρτηση Map εκτελείται σε κάθε χρήστη και χρησιμοποιείται:

- για την παραγωγή των δεδομένων (ADU) όπως περιγράψαμε παραπάνω (χρησιμοποιώντας τους ενσωματωμένους αισθητήρες της συσκευής)
- για την αποστολή τους στον κεντρικό κόμβο.

Μόλις παραχθούν τα δεδομένα, το ID της πλειάδας και του χρήστη, η γεωγραφική περιοχή

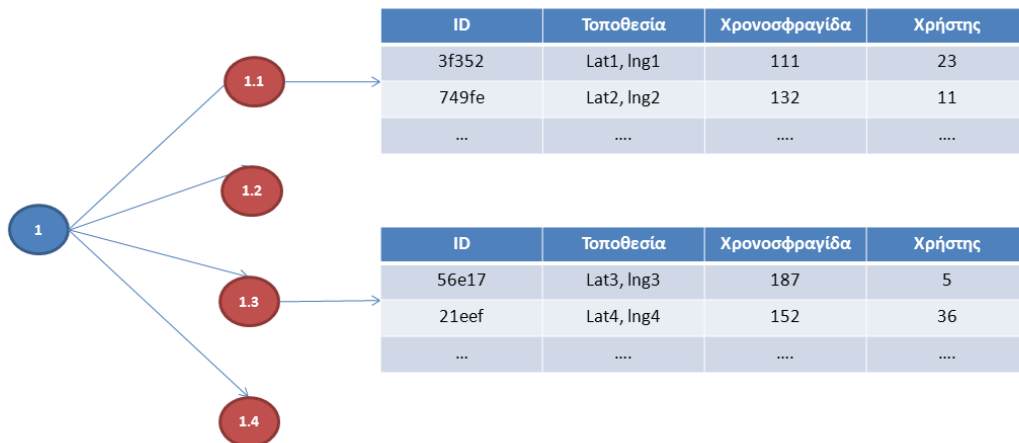
των δεδομένων και η χρονοσφραγίδα μεταφέρονται σε ένα κεντρικό κόμβο σαν έξοδος της Map και σαν είσοδος της Reduce συνάρτησης ώστε να αποθηκευτούν, ως <ID, id_χρήστη, latitude, longitude, χρονοσφραγίδα> (αναφορά πλειάδας).

Η Reduce συνάρτηση παίρνει ως είσοδο τις αναφορές πλειάδων των χρηστών που παράχθηκαν για εκείνη τη χρονική περίοδο. Κάθε reduce συνάρτηση παίρνει τη λίστα των αναφορών που αναφέρονται σε μία συγκεκριμένη γεωγραφική περιοχή και αποθηκεύει τις αναφορές αυτές σε μια δενδρική δομή. Η δενδρική δομή αυτή χρησιμοποιείται για να διαιρέσει κάθε περιοχή σε ένα σύνολο υποπεριοχών σε κάθε επίπεδο του δένδρου (Σχήμα 2). Έτσι στα φύλλα του δένδρου υπάρχουν όλες οι υποπεριοχές και για κάθε μία από αυτές χρησιμοποιείται μία ξεχωριστή reduce συνάρτηση για να αποθηκεύσει τις αναφορές.



Σχήμα 2. Δομή

Για την αποθήκευση των αναφορών σε κάθε φύλλο χρησιμοποιούμε μια δομή κατακερματισμού όπου αποθηκεύεται το σύνολο των αναφορών που ανήκουν σε αυτή τη περιοχή (Σχήμα 3). Οι αναφορές που αποθηκεύονται μέσω της Reduce για κάθε ADU είναι τα εξής: <ID, τοποθεσία, χρονοσφραγίδα, id_χρήστη>.



Σχήμα 3. Δομή Κατακερματισμού

Τέλος, τα δεδομένα που αφορούν την εφαρμογή παραμένουν τοπικά στη συσκευή που τα παράγει και μπορούν να ανακτηθούν από αυτή όταν ζητηθούν από κάποιο χρήστη στο MarketBoard, όπως αναφέρεται στην επόμενη παράγραφο.

3.4 Ανάκτηση Δεδομένων

Η ανάκτηση των δεδομένων βασίζεται και αυτή στην αρχιτεκτονική MapReduce. Η αναζήτηση των δεδομένων που θα ανακτηθούν γίνεται στους κεντρικούς κόμβους του συστήματος (όχι στα κινητά) με βάση τις αναφορές των δεδομένων.

Για την ανάκτηση χρησιμοποιούμε το σύστημα Hadoop. Για να ανακτήσουμε τα δεδομένα αρχικά πρέπει να λάβουμε μία επερώτηση (query) από κάποιο χρήστη η οποία να προσδιορίζει το σύνολο των πλειάδων για τα οποία θέλει να λάβει δεδομένα (πχ τη γεωγραφική περιοχή και το χρονικό διάστημα για το οποίο επιθυμεί κάποιος να λάβει δεδομένα).

Μόλις το σύστημα λάβει ένα τέτοιο επερώτημα ξεκινάει τη Map συνάρτηση η οποία χρησιμοποιεί ένα σύνολο από workers («εργάτες») για να αναγνωρίσουν το σύνολο των ADU που ανήκουν στο εύρος του επερωτήματος τόσο χρονικά όσο και χωρικά. Έτσι οι workers αναζητούν στα υποσύνολα του δένδρου (φύλλα) που ορίσαμε προηγουμένως τις αναφορές που αφορούν τη γεωγραφική περιοχή που αναζητάει ο χρήστης (πχ υποσύνολα 1.1, 1.2, 2.1).

- Κάθε worker εξετάζει ένα ή περισσότερα υποσύνολα του δένδρου (φύλλα), χωρίς όμως να υπάρχουν επικαλύψεις με άλλους εργάτες
- Σε αυτά αναζητάει τις αναφορές των ADU τα οποία ανήκουν στο χρονικό διάστημα και στην περιοχή του επερωτήματος.
- Για κάθε τέτοιο ADU επικοινωνεί με τη συσκευή του χρήστη που διατηρεί το ADU, εφόσον η συσκευή είναι ενεργή και του ζητάει τις αντίστοιχες πλειάδες.
- Κάθε χρήστης που λαμβάνει μία τέτοια αίτηση, προωθεί τα δεδομένα του στον κεντρικό κόμβο.

Τα δεδομένα των χρηστών έρχονται σαν είσοδος στην Reduce συνάρτηση που είναι υπεύθυνη ώστε να ανακτήσει το σύνολο των δεδομένων από όλους τους χρήστες και στη συνέχεια να τα ενώσει ώστε να μπορέσει να τα επιστρέψει στον χρήστη που έκανε την επερώτηση σαν αποτέλεσμα.

3.5 Ενδεικτικό σενάριο χρήσης

Έστω ότι έχουμε ένα σενάριο εφαρμογής για παρακολούθηση της κίνησης στους δρόμους. Οι χρήστες παράγουν δεδομένα της μορφής $\langle \text{carID}, \text{χρονοσφραγίδα}, \text{latitude}, \text{longitude}, \text{speed} \rangle$ (Berkeley's Mobile Millennium Dataset) ενώ από αυτά τα δεδομένα είναι εύκολο να εξάγουμε το ID της πλειάδας χρησιμοποιώντας hashing.

Οι χρήστες παράγουν τα δεδομένα από τους αισθητήρες της συσκευής – στη συγκεκριμένη εφαρμογή χρησιμοποιείται μόνο ο αισθητήρας GPS – και στέλνουν τις αναφορές των δεδομένων όπως αναφέρθηκε παραπάνω στους reduce κόμβους. Κάθε reduce κόμβος που είναι υπεύθυνος για τη συγκεκριμένη γεωγραφική περιοχή αποθηκεύει τα δεδομένα όπως αναφέρθηκε στο 3.3. Έστω ότι κάποιος χρήστης αναζητάει τη μέση ταχύτητα της περιοχής x για τη χρονική περίοδο t_1-t_2 . Θα κάνει το αντίστοιχο επερώτημα και το σύστημα θα ξεκινήσει τη διαδικασία ανάκτησης δεδομένων

- Θα αναθέσει σε κάθε worker ένα φύλλο από τα φύλλα που ανήκουν στην περιοχή x
- Κάθε worker θα βρει το υποσύνολο των αναφορών στο φύλλο που εξετάζει που ανήκουν στη συγκεκριμένη χρονική περίοδο (και περιοχή)

- Θα ανακτήσει τα δεδομένα από τα κινητά που κατέχουν τα συγκεκριμένα δεδομένα

Στο συγκεκριμένο ερώτημα δεν είναι απαραίτητο να ανακτηθούν όλα τα δεδομένα αλλά μπορούμε να εξάγουμε ένα υποσύνολο από εκείνα που ανταποκρίνονται στα κριτήρια αναζήτησης και στους περιορισμούς που θέτει ο χρήστης.

Τέλος τα δεδομένα αυτά θα επιστραφούν στον χρήστη που τα ζήτησε ώστε να υπολογίσει τη μέση ταχύτητα από τις ταχύτητες των πλειάδων που του επιστράφηκαν

3.5.1 Ενσωμάτωση Μηχανισμών Αγοράς

Στο σύστημα αυτό μπορούν να προταθούν διάφοροι μηχανισμοί αγοράς (χρέωση, δημοπρασίες κτλ.) ώστε οι χρήστες να επωφελούνται από την προσφορά δεδομένων. Με αυτό τον τρόπο μπορούμε να τους δώσουμε κίνητρα ώστε να συμμετέχουν στο σύστημα και να παράγουν δεδομένα.

Οι συγκεκριμένοι μηχανισμοί θα συζητηθούν εκτενέστερα στο Κεφάλαιο 4, και αποτελούν ένα από τα βασικά αντικείμενα της έρευνας του INCEPTION. Παρακάτω παραθέτουμε ένα παράδειγμα ενός απλού μηχανισμού αγοράς βάσει χρέωσης και πως σχετίζεται με τη δομή MarketBoard και ενδιάμεσου λογισμικού.

Στην πιο απλή μορφή όλα τα δεδομένα θα έχουν την ίδια τιμή. Έτσι θα μπορούσαμε να θεωρήσουμε ότι κάθε χρήστης που κάνει μια επερώτηση θέτει και ένα μέγιστο όριο πληρωμής B , οπότε επιλέγονται από τους χρήστες k ADU ώστε $k \cdot price < B$ και επιστρέφονται στο χρήστη.

Πρακτικά όμως κάθε ADU έχει ξεχωριστό Quality Of Information που μπορεί να εξαρτάται από παράγοντες όπως η πυκνότητα αισθητήρων στην περιοχή, τεχνολογική τους αξιοπιστία κτλ. Επίσης μπορεί οι χρήστες να θέλουν να πληρώνονται:

- σε σχέση με το πόσο μεγάλο είναι το QoI της πλειάδας
- μπορεί να έχουν επιπρόσθετα κόστη (κατανάλωση μπαταρίας) και έτσι κάθε χρήστης να θέτει διαφορετική τιμή.

Σε αυτή τη περίπτωση θα θέλαμε να επιλέξουμε τα δεδομένα που παρέχουν την πιο σημαντική/ποιοτική πληροφορία (QoI) σε σχέση με τη τιμή των πλειάδων που θέτει ο χρήστης ώστε να μην ξεπεράσουμε την τιμή B . Αυτό αντιστοιχεί στο γνωστό θεωρητικό πρόβλημα υψηλής πολυπλοκότητας «Knapsack problem», για τη προσεγγιστική επίλυση του οποίου μπορεί να εφαρμοσθεί κάποιος ευρετικός αλγόριθμος.

Επίσης σε αυτούς τους μηχανισμούς θα πρέπει να λάβουμε υπ' όψιν και τους περιορισμούς του συστήματος. Για παράδειγμα μπορεί κάποιος να ζητήσει όλα τα δεδομένα στο σύστημα χωρίς να τον ενδιαφέρει η τιμή B , όμως για να γίνει αυτό αποδοτικά θα πρέπει το σύστημα να μπορεί ανταπεξέλθει αποδοτικά αφού θα υπάρχουν:

- Περιορισμοί επεξεργασίας
- Περιορισμοί επικοινωνίας
- Τα κινητά θα επιβαρυνθούν πολύ και ενώ θα ανταμειφθούν χρηματικά οι χρήστες μπορεί οι συσκευές τους να «μπλοκάρουν» για κάποιο χρονικό διάστημα και να μην μπορούν να τις χρησιμοποιήσουν

4. Παρουσίαση της Υπηρεσίας Αγοράς (Market service) και των μηχανισμών κινήτρων

Στο παρόν κεφάλαιο γίνεται η παρουσίαση της Υπηρεσίας Αγοράς (Market service) και των μηχανισμών κινήτρων σύμφωνα με τα σενάρια επίδειξης του συστήματος, τα οποία επίσης αναπτύσσονται στο παρόν κεφάλαιο.

Οι παρακάτω ορισμοί θα χρησιμοποιηθούν για την περιγραφή των μηχανισμών της Υπηρεσίας Αγοράς

- Έστω σύνολο $G = \{1, 2, 3, \dots, n\}$ που αντιπροσωπεύει το διαθέσιμο σύνολο των διαφόρων ομάδων (τύπων) των αισθητήρων (π.χ. το GPS, επιταχυνσιόμετρα, φως, υγρασία, CO2 κλπ). Κάθε ομάδα αποτελείται από έναν αριθμό όμοιων αισθητήρων τα δεδομένα των οποίων ως ψηφιακά δεν καταναλώνονται όταν χορηγούνται σε πλειοδότη, δηλαδή μπορούν να διατεθούν σε πολλούς πλειοδότες.
- Έστω σύνολο $I = \{1, 2, 3, \dots, m\}$ που αντιπροσωπεύει το σύνολο των παρόχων (ή προμηθευτών) των δεδομένων για κάθε αισθητήρα, με εγγυημένη ανωνυμία. Η τιμή m είναι μια ένδειξη της μέγιστης δυνατής προσφοράς δεδομένων σε κάθε t .
- Έστω σύνολο $J = \{1, 2, 3, \dots, k\}$ το σύνολο των πιθανών αγοραστών των δεδομένων του αισθητήρα (οι πλειοδότες). Κάθε υποψήφιος αγοραστής j από το σύνολο J , απαιτεί να αγοράσει πλειάδες από ένα υποσύνολο $S_j \subseteq G$. Επίσης, το k υποδηλώνει το μέγιστο αριθμό πιθανών υποψηφίων σε κάθε t , και αποτελεί μια ένδειξη της ζήτησης. Σημειωτέον ότι ένας χρήστης μπορεί άλλοτε να έχει τον ρόλο παρόχου και άλλοτε τον ρόλο αγοραστή.
- η μεταβλητή s_{ig} υποδηλώνει τη διαθεσιμότητα (0 ή 1) ενός αισθητήρα που ανήκει στην ομάδα G (π.χ. σε $g = 1$) από τον προμηθευτή i
- η παράμετρος $x(s_{ig})$ συμβολίζει την πραγματική αξία του αισθητήρα s_{ig}
- $r(s_{ig})$ ή r_{ig} είναι η ελάχιστη τιμή (reserve price) για τον πάροχο i (ελάχιστο ποσό που είναι διατεθειμένος να λάβει) για την παροχή ενός αισθητήρα που ανήκει σε μια ομάδα. Για λόγους απλότητας θα το συμβολίζουμε ως r_{ig} . Για παράδειγμα, R_{11} είναι η ελάχιστη τιμή για τον αισθητήρα S_{11} δηλαδή που παρέχεται από τον πάροχο 1 και ανήκει στην ομάδα 1 (δηλαδή $g = 1$).
- Ως S_j θα συμβολίζουμε ένα υποσύνολο από την ομάδα των διαθέσιμων αισθητήρων $S_j \subseteq G$
- Ο πίνακας που ακολουθεί δείχνει τη δομή όπου τηρείται η διαθεσιμότητα των δεδομένων σε ένα συγκεκριμένο χρονικό διάστημα t και μπορεί να αναπαρασταθεί ως ένας πίνακας 2 διαστάσεων $n \times m$.

	Groups of available sensors (G)			
Providers	1	2	...	n
1	S_{11}	S_{12}	...	S_{1n}

2	S ₂₁	S ₂₂	...	S _{2n}
3	S ₃₁	S ₁₂	...	S _{3n}
...
m	S _{m1}	S _{m2}	...	S _{mn}

- B_j = η προσφορά που γίνεται από τον πλειοδότη j για ένα υποσύνολο S_j
- n_j = αριθμός των πλειάδων που ζητούνται από τον πλειοδότη j (για το σύνολο S_j)
- p_j = τιμή (χρέωση) που πρέπει να καταβληθεί για κάθε πλειάδα ή το \vec{p}_j διάνυσμα των τιμών (θα συζητηθεί αργότερα).

4.1 Πιθανά Σενάρια Επίδειξης του Συστήματος

Υπάρχουν πολλές πιθανές επιλογές για τα σενάρια υλοποίησης της υπηρεσίας αγοράς και των μηχανισμών αυτής. Για να επιλεγθούν τα πιο ενδεδειγμένα πρέπει πρώτα να κατηγοριοποιηθούν οι διάφορες επιλογές και να αναλυθεί η καταλληλότητα τους όπως και οι τεχνικές επιπτώσεις επιλογής αυτών.

Η σχετική κατηγοριοποίηση παρουσιάζεται πιο κάτω με βάση τέσσερα κριτήρια.

A. Τύπος εφαρμογής

A1. Εφαρμογές που απαιτούν σε πραγματικό χρόνο ενημέρωση (streaming) των δεδομένων, π.χ. πλοήγησης, στάθμευσης κλπ

A2. Εφαρμογές που σχετίζονται με μία συγκεκριμένη τοποθεσία και είναι χρονικά σταθερές π.χ. περιβαλλοντολογικές

B. Περιορισμούς στις πλειάδες που ζητά ο πλειοδότης

B1. Ο πλειοδότης j ζητά n_j πλειάδες, αλλά είναι διατεθειμένος να παρεί και λιγότερες από αυτές (δηλαδή η τιμή n_j είναι ο μέγιστος αριθμός από πλειάδες που ενδιαφέρεται να πάρει). Σε αυτή την περίπτωση υποβάλλει

B11. Μια τιμή ανά πλειάδα (ίδια χρησιμότητα για όλες)

B12. Διάνυσμα τιμών για τις διάφορες πλειάδες

B2. Ο πλειοδότης j ζητά n_j πλειάδες και είναι ικανοποιημένος μόνο με τον ακριβή αριθμό των πλειάδων και όχι λιγότερες/περισσότερες. Ως εκ τούτου, υποβάλλει μία προσφορά που προσδιορίζει το συνολικό τίμημα που μπορεί να καταβληθεί.

Γ. Τιμές Παρόχων

Γ1. Οι πάροχοι ορίζουν μία τιμή για κάθε αισθητήρα ξεχωριστά ή οποία καταβάλλεται μία φορά (ασχέτως από πόσους θα ζητηθεί ο συγκεκριμένος αισθητήρας). Σε αυτή την περίπτωση ο επιμερισμός του κόστους μεταξύ των πλειοδοτών

Γ11. Επιτρέπεται ή

Γ12. Δεν επιτρέπεται

Γ2. Έχει οριστεί μία τιμή για όλους τους αισθητήρες, δηλαδή ανεξάρτητα από τον αριθμό των αισθητήρων που θα χρησιμοποιηθούν σε κάθε στιγμή η τιμολόγηση είναι η ίδια.
Γ3. Ορίζεται μια αρχική τιμή (set-off) τιμή που αντιπροσωπεύει το κόστος για την ενεργοποίηση της συσκευής και επιπλέον οι τιμές για κάθε αισθητήρα που θα προσφέρει ο πάροχος

Δ. Συνδυαστικές προσφορές από τους πλειοδότες

Δ1. Δεν επιτρέπονται. Οι πλειοδότες υποβάλλουν προσφορές για ένα υποσύνολο μόνο.
Δ2. Επιτρέπονται. Οι πλειοδότες μπορούν να υποβάλουν προσφορές για συνδυασμούς αισθητήρων (διάφορα υποσυνόλα).

Βασικές Παραδοχές (Ισχύουν για όλες τις περιπτώσεις)

- Οι προσφορές των πλειοδοτών είναι ειλικρινείς, δηλ. οι πλειοδότες δεν αποκρύπτουν την αξία που έχουν για αυτούς οι αισθητήρες.

4.2 Επιλογή και περιγραφή του βασικού σεναρίου

Με βάση τις απαιτήσεις του έργου και των σεναρίων αξιολόγησης επιλέχθηκαν τα σενάρια τα οποία ανήκουν στις συγκεκριμένες κατηγορίες:

- A2. Εφαρμογές που σχετίζονται με μία συγκεκριμένη τοποθεσία και είναι χρονικά σταθερές π.χ. περιβαλλοντολογικές

B11. Ο πλειοδότης ζητά $n_j = \max$ πλειάδες, αλλά είναι διατεθειμένος να πάρει και λιγότερες από αυτές. Σε αυτή την περίπτωση υποβάλλει μια τιμή ανά πλειάδα (ίδια χρησιμότητα για όλες)

- Γ1. Οι πάροχοι ορίζουν μία τιμή για κάθε αισθητήρα ξεχωριστά, η οποία καταβάλλεται μία φορά (ασχέτως από πόσους θα ζητηθεί ο συγκεκριμένος αισθητήρας). Σε αυτή την περίπτωση ο επιμερισμός του κόστους μεταξύ των πλειοδοτών επιτρέπεται (Γ11)
- Δ1. Οι συνδυαστικές προσφορές από τους πλειοδότες δεν επιτρέπονται. Οι πλειοδότες υποβάλλουν προσφορές για ένα υποσύνολο μόνο

Αν και δεν είναι το πιο γενικό σενάριο που μπορεί να αναπτυχθεί, το ανωτέρω βασικό σενάριο περιλαμβάνει πολλά ενδιαφέροντα χαρακτηριστικά τα οποία χρήζουν αναλύσεως και αντιμετώπισης με ειδικούς αλγορίθμους μηχανισμών αγοράς.

Ο στόχος του βασικού σεναρίου είναι η δημιουργία μιας βιώσιμης και αποτελεσματικής αγοράς (προσέλκυση προμηθευτών και αγοραστών και ικανοποίηση τους μέσω της

αντιστοίχισης της προσφοράς με τη ζήτηση με τον καλύτερο δυνατό τρόπο) και όχι η μεγιστοποίηση των οικονομικών οφελών του ιδιοκτήτη της αγοράς.

Με βάση την παραπάνω επιλογή οι ρόλοι των βασικών παικτών του σεναρίου είναι οι εξής:

Πάροχοι δεδομένων αισθητήρων

- Παρέχονται πλειάδες αισθητήρων. Για το βασικό σενάριο θα λάβουμε υπόψη μας την διάθεση λίγων τύπων αισθητήρων. Κάθε πάροχος αναφέρει ποιούς από τους αισθητήρες (ή όλους) θα θέσει στη διάθεση στην αγορά και μια ελάχιστη τιμή για κάθε αισθητήρα.
- Μόλις πληρωθεί αυτή η ελάχιστη τιμή (ενδεχομένως) συλλογικά από τους υποψηφίους, τα δεδομένα διαβιβάζονται και ο προμηθευτής πληρώνεται αυτή την τιμή (μία φορά), ανεξάρτητα από το σε πόσους υποψηφίους θα είναι διαθέσιμα τα δεδομένα.

Αγοραστές (Πλειοδότες)

- Οι πλειοδότες υποβάλουν προσφορές για την απόκτηση συγκεκριμένου αριθμού πλειάδων, αλλά είναι επίσης πρόθυμοι να αγοράσουν λιγότερες από αυτό τον αριθμό.
- Η προσφορά που υποβάλλεται (ποσό που πρέπει να καταβληθεί) αντιστοιχεί στην αγορά μιας πλειάδας. Ο μέγιστος αριθμός των πλειάδων επίσης δηλώνεται. Ωστόσο, η χρησιμότητα της αγοράς για την καθεμιά από αυτές είναι η ίδια.
- Αρχικά υποθέτουμε ομοιογενή ζήτηση των πλειάδων, με την έννοια ότι όλοι απαιτούν τον ίδιο τύπο πλειάδων, αλλά ο καθένας προσφέρει μια διαφορετική τιμή ανά πλειάδα και ζητά ένα διαφορετικό μέγιστο αριθμό των πλειάδων. Μετέπειτα, θα αφαιρέσουμε την υπόθεση αυτή και θα μελετήσουμε σχετικούς μηχανισμούς αγοράς και θα συγκρίνουμε τα αποτελέσματα με την παρούσα περίπτωση.
- Το κόστος για την αγορά των πλειάδων μπορεί να μοιραστεί μεταξύ των διαγωνιζομένων. Παρ' όλα αυτά, ο μηχανισμός της αγοράς αποφασίζει πώς θα πρέπει να μοιραστεί το κόστος χωρίς να υπάρχει αλληλεπίδραση μεταξύ τους
- Υποθέτουμε ότι οι πλειοδότες είναι ειλικρινείς κατά την υποβολή προσφορών τους.

Ακόμη και σε αυτό το απλό σενάριο με συγκεκριμένες παραδοχές (π.χ. λίγους αισθητήρες, τιμή ανά αισθητήρα, κλπ) και περιορισμούς (π.χ. ομοιογενής ζήτηση) μπορούμε να κάνουμε ενδιαφέρουσες παρατηρήσεις και να αναλύσουμε προβλήματα. Η άρση ορισμένων περιορισμών ή/και των παραδοχών κάνει το πρόβλημα της εφικτότητας (feasibility) και της κατανομής των νικητών (allocation) δύσκολα να αντιμετωπιστούν. Στοιχεύουμε στην εφαρμογή και τη σύγκριση των επιδόσεων των διάφορων μηχανισμών κατανομής και των συναφών κανόνων τιμολόγησης (συστήματα χρέωσης) με βάση ένα συγκεκριμένο σύνολο κριτηρίων.

4.3 Παράδειγμα υλοποίησης μηχανισμού της Υπηρεσίας Αγοράς για το Βασικό Σενάριο

Αρχίζουμε με την περίπτωση όπου όλοι οι πλειοδότες έχουν τον ίδιο τύπο ζήτησης, δηλαδή έχουμε ομοιόμορφη ζήτηση από τους αγοραστές. Όλοι ζητούν τον ίδιο τύπο πλειάδων, ο οποίος αποτελείται από μια ομάδα αισθητήρων, π.χ. θερμοκρασίας και CO₂. Στο παράδειγμα μας ας υποθέσουμε ότι όλοι θέλουν πλειάδες από 2 τύπους αισθητήρων. Ακόμη και αν αυτό φαίνεται να είναι μια απλή υπόθεση, δεν είναι εκ των προτέρων σαφές ποιος μηχανισμός θα λειτουργήσει καλύτερα, καθώς υπάρχουν αρκετές παραλλαγές που μπορεί κανείς να εξετάσει. Αυτό που διαφοροποιεί τους αγοραστές σε αυτή τη ρύθμιση, είναι ο μέγιστος αριθμός των πλειάδων που θέλουν (που όπως έχει ήδη αναφερθεί προηγουμένως συμβολίζεται με n_j για αγοραστή j) και η τιμή ανά πλειάδα που είναι πρόθυμοι να πληρώσουν (δηλαδή, p_j για τον αγοραστή j). Αυτό σημαίνει ότι κάθε αγοραστής έχει συνολικό προϋπολογισμό $B_j = n_j p_j$, που είναι πρόθυμος να δαπανήσει το πολύ.

Οι μηχανισμοί που θα εξετάσουμε λειτουργούν σε γύρους (rounds) και χρεώνουν ομοιόμορφα για κάθε πλειάδα που διαθέτουν, αρχής γενομένης από την φθηνότερη. Σε κάθε γύρο, όπως εξηγείται παρακάτω, ο μηχανισμός εξετάζει την τρέχουσα τιμή ανά πλειάδα που κάθε πλειοδότης είναι πρόθυμος να πληρώσει. Ακόμα κι αν οι αγοραστές καθορίσουν μια τιμή ανά πλειάδα στην αρχή, ως μέρος της προσφοράς τους, αυτό μπορεί να προσαρμοσθεί λόγω της εξοικονόμησης που θα προκύψει εάν πληρώνουν λιγότερο για τις πλειάδες στους προηγούμενους γύρους (και το υπόλοιπο από τα χρήματα μπορεί να «μεταφερθεί» στον επόμενο γύρο), ή επειδή μπορεί ακόμη και να είναι σύμφωνοι με το να ξοδέψουν ολόκληρο τον προϋπολογισμό τους για την αγορά λιγότερων πλειάδων. Η συζήτηση αυτή οδηγεί σε εναλλακτικά σενάρια, ανάλογα με το πώς αντιλαμβανόμαστε την ανώτατη τιμή που ένας αγοραστής είναι πρόθυμος να πληρώσει σε κάθε γύρο για μία πλειάδα.

Αρχικά ο μηχανισμός συλλέγει το σύνολο των προσφορών και της ζήτησης περιοδικά μετά από ένα προκαθορισμένο χρονικό διάστημα, π.χ. 30 λεπτά. Με βάση αυτές τις πληροφορίες, το σύστημα δημιουργεί ένα σύνολο από πίνακες ή παρόμοια δομή, μια για κάθε αισθητήρα για να βοηθήσει στην εφαρμογή των αλγορίθμων και της αξιολόγησης. Επίσης παράγει τις ακόλουθες πληροφορίες (ασχέτως αν χρειαστούν όλες σε κάθε περίπτωση):

- Τον συνολικό αριθμό των παρόχων που προσφέρουν αυτόν τον αισθητήρα - μια ένδειξη της προσφοράς
- Τις τιμές προσφοράς για αυτόν τον αισθητήρα σε αύξουσα σειρά (αρχίζοντας από την φθηνότερη) και το σχετικό πάροχο που προσφέρει κάθε μία από αυτές
- Τον συνολικό αριθμό των υποψηφίων αγοραστών για αυτόν τον αισθητήρα, είτε ζητείται ως αυτόνομα ή ως μέρος μιας πλειάδας - μια ένδειξη της ζήτησης
- Τον συνολικό αριθμό των ζητούμενων πλειάδων για αυτόν τον αισθητήρα - μια ένδειξη της ζήτησης. Αυτό είναι διαφορετικό από το προηγούμενο, καθώς ένας υποψήφιος αγοραστής μπορεί να ζητήσει πολλαπλές φορές τον ίδιο αισθητήρα (από διαφορετικούς παρόχους)
- Την συνολική προσφορά χρημάτων για κάθε αισθητήρα, προσθέτοντας όλες τις προσφορές

DEMAND (BIDS)				
Bidder	S ₁	S ₂	Tuples wanted	Price (bid) per tuple
B ₁	1	1	5	2
B ₂	1	1	2	2
B ₃	1	1	1	3
B ₄	1	1	3	3
B ₅	1	1	2	1
B ₆	1	1	2	4
B ₇	1	1	1	7

SUPPLY – Reserve prices	
g ₁	g ₂
i ₂ = 6	i ₄ = 2
i ₃ = 8	i ₁ = 4
i ₅ = 12	i ₃ = 8
	i ₅ = 14

Πίνακας 1. Παράδειγμα πινάκων ζήτησης και προσφοράς για το σενάριο A1

Περιγραφή των βημάτων του μηχανισμού

1. Αρχικά εξετάζουμε την πρώτη και φθηνότερη πλειάδα συνδυάζοντας το φθηνότερο αισθητήρα από κάθε τύπο των αισθητήρων που ζητούνται στην πλειάδα. Έχουμε ελέγξει πόσοι από τους υποψηφίους έχουν ζητήσει αυτή την πλειάδα (δύο τύπων αισθητήρων 1 και 2 στο παράδειγμά μας), ανεξάρτητα από το εάν έχουν ζητήσει μία ή περισσότερες από αυτές.
2. Πρώτα προσπαθούμε να μοιράσουμε το κόστος εξ ίσου μεταξύ εκείνων που την έχουν ζητήσει. Π.χ. αν υποθέσουμε ότι το 8 είναι η ζητούμενη τιμή (το ποσό για τους δύο αισθητήρες στο παράδειγμά μας), διαιρούμε τον αριθμό με το συνολικό αριθμό των πλειοδοτών "k" που θέλουν αυτή την πλειάδα, π.χ. 7 στην περίπτωση μας και ελέγχουμε πόσοι από αυτούς μπορούν να πληρώσουν αυτό το ποσό. Σε αυτό το σημείο έχουμε διαφορετικές επιλογές για τους υποψήφιους αγοραστές που δεν μπορούν να καλύψουν αυτή την τιμή. Αυτός είναι ο 1ος βαθμός ελευθερίας στο σενάριο μας με βάση τα προγράμματα επιμερισμού του κόστους.

v1. Χρεώνουμε εξίσου όλους τους πλειοδότες που μπορούν να πληρώσουν τις τιμές που έχουν ζητήσει οι δύο πάροχοι με $(rs_{i1}+rs_{i2})/k$ και αποκλείουμε τους υποψήφιους αγοραστές που δεν μπορούν να καλύψουν αυτή τη χρέωση από την κατανομή μας και, συνεπώς, και κατά τους επόμενους γύρους (για την κατανομή της 2ης πλειάδας κ.ο.κ.). Στη συνέχεια, επιμερίζουμε το κόστος του αισθητήρα, με βάση το νέο n , το οποίο έχει μειωθεί αυξάνοντας έτσι το κόστος του αισθητήρα ανά πλειοδότη, μέχρι να βρούμε τον αριθμό των υποψηφίων που μπορούν να καλύψουν το κόστος του αισθητήρα. Ωστόσο, υπάρχει πάντα η πιθανότητα επειδή η τιμή που καλούνται να πληρώσουν σταδιακά αυξάνεται και οι υποψήφιοι μειώνονται να φτάσουμε σε ένα σημείο που δεν υπάρχουν υποψήφιοι να πληρώσουν για την πρώτη πλειάδα.

v2. Ίδια προσέγγιση με την προηγούμενη περίπτωση, αλλά δεν αποκλείουμε τους υποψήφιους αγοραστές που δεν μπορούν να καλύψουν το συγκεκριμένο

ποσό. Αντί αυτού, τους χρεώνουμε με το ποσό που μπορούν να πληρώσουν (με βάση την προσφορά τους) και το υπόλοιπο ποσό που απαιτείται για να καλυφθεί το κόστος της πλειάδας το μοιράζουμε μεταξύ των άλλων υποψηφίων είτε εξίσου είτε αναλογικά με βάση την προσφορά που είχαν κάνει. Εάν κάποιος άλλος υποψήφιος στη συνέχεια δεν μπορεί να πληρώσει αυτό το κόστος ακολουθούμε την ίδια διαδικασία με τον καταμερισμό σε αυτούς που μπορούν να πληρώσουν και τον καταμερισμό του υπόλοιπου ποσού στους υπολειπόμενους πλειοδότες.

3. Συνεχίζουμε με την επόμενη φθηνότερη πλειάδα επαναλαμβάνοντας τα προηγούμενα βήματα μέχρι

- α) Την ικανοποίηση της ζήτησης: όλοι οι υποψήφιοι λαμβάνουν τον αριθμό των πλειάδων που ζήτησαν
- β) Την έλλειψη ζήτησης: κανένας υποψήφιος δεν μπορεί να πληρώσει την επόμενη πλειάδα
- γ) Την έλλειψη προσφοράς: χρειαζόμαστε περισσότερες πλειάδες να ικανοποιήσουμε την ζήτηση

4. Ωστόσο, υπάρχει ένας 2ος βαθμός ελευθερίας όσον αφορά τις επιλογές που μπορούμε να κάνουμε ανάλογα με το πώς λαμβάνουμε υπόψη την ανώτατη τιμή που ένας αγοραστής είναι πρόθυμος να πληρώσει σε κάθε γύρο για μία πλειάδα και το πώς αντιμετωπίζουμε το υπόλοιπο ποσό μετά από κάθε κατανομή

v1. Ο συνολικός προϋπολογισμός που απομένει από κάθε συμμετέχοντα (αρχικά ίσος με B_j για κάθε j) είναι διαθέσιμος για να αγοράσει κάθε πλειάδα.

v2. Η τιμή ανά πλειάδα, p_j , που υποβλήθηκε αρχικά από τον αγοραστή, είναι διαθέσιμη για την αγορά μόνο αυτής της συγκεκριμένης πλειάδας.

v3. Για την πρώτη πλειάδα, χρησιμοποιείται η υποβληθείσα τιμή ανά πλειάδα. Αν η χρέωση είναι μικρότερη, τα υπολειπόμενα χρήματα μεταφέρονται στον 2ο γύρο, αυξάνοντας τα διαθέσιμα χρήματα ανά πλειάδα για το συγκεκριμένο γύρο. Σε γενικές γραμμές οποιαδήποτε υπόλοιπα χρημάτων από τους προηγούμενους γύρους, προστίθενται στην p_j για τον προσδιορισμό της τιμής ανά πλειάδα του επόμενου γύρου.

5. Σενάρια αξιολόγησης και κριτήρια επιτυχίας

Τα σενάρια και τα κριτήρια αξιολόγησης του υπό μελέτη συστήματος θα κινηθούν σε τρεις άξονες:

- 1. Αξιοπιστία.** Το σύστημα θα πρέπει να τρέχει χωρίς σφάλματα όταν οι συσκευές παράγουν και αποθηκεύουν τα δεδομένα καθώς και όταν ανακτούν δεδομένα. Θεωρούμε ότι για να πληρούται το κριτήριο θα πρέπει να τρέχει ορθά σε τουλάχιστον 2 συσκευές. Επίσης το σύστημα θα πρέπει να είναι σε θέση να ανακτά τα δεδομένα εφόσον είναι διαθέσιμα στο MarketBoard και να ακολουθεί πιστά το μηχανισμό κινήτρων που έχει επιλεγεί.
- 2. Κλιμακωσιμότητα.** Το σύστημα θα πρέπει να δουλεύει το ίδιο αποδοτικά όσο αυξάνεται το σύνολο των χρηστών και των δεδομένων. Επίσης η πολυπλοκότητα των μηχανισμών θα είναι τέτοια ώστε η επιλογή των δεδομένων που θα πουληθούν στους πλειοδότες να γίνεται σε αποδεκτούς χρόνους ανεξαρτήτως του μεγέθους των πλειοδοτών και των διαθέσιμων δεδομένων. Παρόμοια θα πρέπει το σύστημα να δουλεύει αποδοτικά ακόμα και όταν οι προσφορές για αγορά δεδομένων ανα χρονική στιγμή αυξάνονται.
- 3. Αποδοτικότητα.** Θα πρέπει το σύστημα να μην επιβαρύνει αρκετά τις συσκευές των χρηστών, ώστε να μην τους αποτρέπουμε να χρησιμοποιούν το σύστημα μας για να προσφέρουν δεδομένα. Επίσης, θα πρέπει τα δεδομένα να μπορούν να ανακτηθούν σε χρόνο ώστε να είναι ανεκτός από τους χρήστες που τα ζητάνε και θα βελτιστοποιεί τους στόχους του συστήματος.

Επίσης η επιτυχία της υπηρεσίας αγοράς θα αξιολογηθεί μετρώντας την απόδοση των προτεινόμενων μηχανισμών με βάση και τα παρακάτω κριτήρια:

- 1. Καταμερισμό κόστους (Cost sharing):** Οι μηχανισμοί και οι αλγόριθμοι υλοποίησης τους θα πρέπει να εξασφαλίζουν ένα δίκαιο καταμερισμό κόστους ώστε περισσότεροι πλειοδότες να λαμβάνουν τις πλειάδες που έχουν ζητήσει με το λιγότερο δυνατό κόστος.
- 2. Αποδοτική επίλυση του προβλήματος κατανομής (allocation problem):** Οι μηχανισμοί θα πρέπει να εξασφαλίζουν ότι όσο το δυνατόν περισσότεροι πλειοδότες θα λαμβάνουν τις πλειάδες που ζήτησαν και παράλληλα περισσότεροι πάροχοι θα ικανοποιούνται (πληρώνονται τις τιμές που ζήτησαν). Επίσης αυτό πρέπει να γίνεται στον λιγότερο δυνατό χρόνο (π.χ. λίγους γύρους δημοπρασιών).
- 3. Social Welfare (Κοινωνική Ευημερία):** Σαν όρος των οικονομικών η κοινωνική ευημερία ορίζεται ως η συνολική χρησιμότητα την οποία απολαμβάνουν οι χρήστες (στην παρούσα περίπτωση προμηθευτές και αγοραστές δεδομένων) από την εκχώρηση τους μέσω κάποιου μηχανισμού. Στην παρούσα περίπτωση η εκχώρηση και η χρέωση για δεδομένα καθορίζονται μέσω μηχανισμών αγοράς. Συνεπώς, οι μηχανισμοί που θα μελετηθούν πρέπει να οδηγούν σε υψηλή κοινωνική ευημερία, δηλ. σε ευρεία εκχώρηση σε αγοραστές των προσφερομένων από προμηθευτές δεδομένων.

6. Σύνοψη - Συμπεράσματα

Το παρόν παραδοτέο αποτελεί το πρώτο σημαντικό ορόσημο στην πορεία των εργασιών του έργου INCEPTION προς τον τελικό στόχο του. Συγκεκριμένα το παραδοτέο παρουσιάζει τις απαιτήσεις που καλείται να υποστηρίξει η υπό ανάπτυξη πλατφόρμα (σε οικονομικό, τεχνολογικό και κοινωνικό επίπεδο), την αρχιτεκτονική της πλατφόρμας αυτής, περιλαμβανομένης και της λογικής στην οποία θα βασίζεται η οικονομική βελτιστοποίηση που θα συντελείται στα πλαίσια της αρχιτεκτονικής, και τα σενάρια για την αξιολόγησης της.

Στα πλαίσια αυτά, αρχικά παρουσιάστηκε η σύνοψη της αρχιτεκτονικής μαζί με τις βασικές απαιτήσεις υλοποίησής της, τόσο από τεχνικής απόψεως όσο και από την πλευρά των εφαρμογών. Οι πιο σημαντικές απαιτήσεις αφορούν: την υποστήριξη της υπηρεσίας αγοράς και των μηχανισμών κινήτρων, την υποστήριξη σύγχρονων έξυπνων συσκευών αλλά και την ανωνυμία και ασφάλεια του χρήστη. Για την ικανοποίηση των απαιτήσεων, η αρχιτεκτονική αυτή περιλαμβάνει: α) το Υποσύστημα Οικονομικών Μηχανισμών (Economic Mechanisms Overlay), που περιλαμβάνει και εκτελεί τους μηχανισμούς αγοράς και διαχειρίζεται τις περιγραφές των διατιθεμένων από χρήστες δεδομένων τα οποία εκχωρούνται μέσω των μηχανισμών αυτών, και β) Ενδιάμεσο Λογισμικό (Connectivity Middleware), που διαχειρίζεται τα ίδια τα διατιθέμενα από χρήστες δεδομένα.

Η αρχιτεκτονική στη συνέχεια αναλύθηκε λεπτομερώς σε συνδυασμό με ένα ενδεικτικό σενάριο χρήσης της. Στο σύστημα μας κάθε χρήστης μπορεί να έχει δύο διακριτούς ρόλους (ή και τους δύο ταυτόχρονα): Να προσφέρει δεδομένα ή να κάνει επερωτήσεις για δεδομένα. Το σύστημα στηρίζεται στην αρχιτεκτονική Map Reduce τόσο για την αποθήκευση όσο και για την ανάκτηση των δεδομένων, και στην προσέγγιση Publish/Subscribe σαν γενική αρχιτεκτονική. Ένα βασικό τμήμα της αρχιτεκτονικής αποτελεί η υπηρεσία αγοράς (market service) και των μηχανισμών κινήτρων. Στο σύστημα θα προταθούν και ενσωματωθούν διάφοροι μηχανισμοί αγοράς (χρέωση, δημοπρασίες κτλ.) ώστε οι χρήστες να επωφελούνται από την προσφορά δεδομένων. Με αυτό τον τρόπο θα παρέχονται τα κατάλληλα κίνητρα στους χρήστες ώστε να συμμετέχουν στο σύστημα και να παράγουν δεδομένα.

Ένα από τα βασικά συμπεράσματα της μελέτης αυτής της αρχιτεκτονικής είναι ότι η αρχιτεκτονική των διάφορων υποσυστημάτων πρέπει να λάβει υπ' όψιν από την αρχική φάση σχεδιασμού τις απαιτήσεις ενσωμάτωσης των οικονομικών μηχανισμών. Για αυτό και αποφασίστηκε οι ομάδες εργασίας που έχουν αναλάβει το σχεδιασμό των τεχνικών υποσυστημάτων όπως του ενδιάμεσου λογισμικού να είναι σε άμεση επικοινωνία με τις ομάδες υλοποίησης των οικονομικών μηχανισμών και να γίνεται σταδιακή ενσωμάτωση αυτών των μηχανισμών στο σύστημα καθώς και τακτικές δοκιμές.

Στη συνέχεια έγινε η περιγραφή των πιθανών σεναρίων επίδειξης του συστήματος και επιλέχθηκε ένα από αυτά ως το βασικό σενάριο για ανάλυση και μετέπειτα υλοποίηση. Τα σενάρια κατηγοριοποιήθηκαν ώστε να είναι σαφής η υποστήριξη τους από τους μηχανισμούς της αγοράς. Η επιλογή του βασικού σεναρίου έγινε ώστε μέσω της ανάλυσης του να εξεταστούν βασικά ερευνητικά θέματα που πραγματεύεται το έργο και να αναπτυχθούν μηχανισμοί κινήτρων οι οποίοι αφενός θα είναι αποδοτικοί για το παρόν σενάριο, αλλά αφετέρου θα μπορούν με κατάλληλες επεκτάσεις να εφαρμοσθούν και σε πιο γενικά σενάρια. Το βασικό σενάριο αφορά εφαρμογές που σχετίζονται με μία συγκεκριμένη τοποθεσία και

χρονική στιγμή. Σε αυτό οι πλειοδότες ζητάν μέγιστο αριθμό από πλειάδες και υποβάλλουν μια τιμή ανά πλειάδα. Οι πάροχοι από την πλευρά τους ορίζουν μία τιμή για κάθε αισθητήρα ξεχωριστά, η οποία καταβάλλεται μία φορά συνολικά από τους πλειοδότες. Σε αυτή την περίπτωση ο επιμερισμός του κόστους μεταξύ των πλειοδοτών επιτρέπεται αλλά όχι οι συνδυαστικές προσφορές, αφού οι πλειοδότες υποβάλλουν προσφορές για ένα υποσύνολο μόνο. Στη πλαίσια αυτά παρουσιάστηκε επίσης ένας μηχανισμός δημοπρασίας της υπηρεσίας αγοράς για το συγκεκριμένο σενάριο.

Ένα από τα βασικά συμπεράσματα της ανάλυσης αυτής είναι ότι υπάρχει μεγάλο εύρος σεναρίων που μπορεί να ορισθεί, βάσει υποθέσεων ως προς τα παρεχόμενα δεδομένα και την ζήτηση γι' αυτά, για τα οποία όμως η ανάλυση του βασικού σεναρίου (το οποίο στηρίζεται σε κάποιες απλουστευτικές υποθέσεις, όπως ήδη αναφέρθηκε ανωτέρω) αναμένεται να είναι αρκετά αντιπροσωπευτική. Επίσης, δε θα πρέπει να αποκλίνει η υλοποίηση των μηχανισμών από τον στόχο του βασικού σεναρίου που είναι η δημιουργία μιας βιώσιμης και αποτελεσματικής αγοράς (προσέλκυση προμηθευτών και αγοραστών και ικανοποίηση τους μέσω της αντιστοίχισης της προσφοράς με τη ζήτηση με κατά το δυνατό τον πιο αποτελεσματικό τρόπο) και όχι η μεγιστοποίηση των οικονομικών οφελών του ιδιοκτήτη της αγοράς.

Τέλος, στο παρόν παραδοτέο περιγράφηκαν και τα βασικά σενάρια αξιολόγησης και τα κριτήρια επιτυχίας του συστήματος, τα οποία κατηγοριοποιήθηκαν σε δύο άξονες:

- A. Κριτήρια αξιολόγησης του συστήματος, όπως αυτά που αφορούν την αξιοπιστία, την κλιμακωσιμότητα και την αποδοτικότητα του.
- B. Κριτήρια αξιολόγησης των οικονομικών μηχανισμών, όπως ο καταμερισμός κόστους, η αποδοτική επίλυση του προβλήματος κατανομής και η κοινωνική ευημερία.

Ως συνέχεια αυτού του παραδοτέου, η εργασία του έργου στο άμεσο μέλλον θα οδηγήσει στην προδιαγραφή (σύμφωνα και με τις απαιτήσεις που οριστήκαν στο παρόν παραδοτέο) του σχεδιασμού της υπηρεσίας αγοράς και του μηχανισμού κινήτρων που θα υλοποιηθεί (Παραδοτέο 3.1), καθώς και του ενδιάμεσου λογισμικού, της δομής MarketBoard και των εφαρμογών που θα χρησιμοποιηθούν για την επίδειξη του έργου (Παραδοτέο 4.1).

7. Αναφορές

- [1] A. Thiagarajan, L. Ravindranath, K. LaCurts, S. Madden, H. Balakrishnan, S. Toledo, and J. Eriksson, "Vtrack: accurate, energy-aware road traffic delay estimation using mobile phones", in SenSys, Berkeley, California, USA, November 2009
- [2] S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, Gahng-seop Ahn, and A. T. Campbell, "Metrosense project: People-centric sensing at scale", in SenSys, Boulder, Colorado, USA, Oct-Nov 2006.
- [3] R. K. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. F. Abdelzaher. "Greengps: a participatory sensing fuel-efficient maps application", In MobiSys, San Francisco, California, USA, June 2010.
- [4] A. Dou, V. Kalogeraki, D. Gunopulos, T. Mielikinen, V. Tuulos, S. Foley, and C. Yu. "Data clustering on a network of mobile smartphones", In SAINT, Munich, Germany, July 2011.
- [5] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda. "Peir, the personal environmental impact report, as a platform for participatory sensing systems research", In MobiSys, pages 55–68, Krakow, Poland, June 2009.
- [6] M. Olson, A. H. Liu, M. Faulkner, and K. M. Chandy. "Rapid detection of rare geospatial events: earthquake warning applications", In DEBS, New York, USA, July 2011.
- [7] A. Okcan and M. Riedewald, "Processing Theta-Joins using MapReduce", In Sigmod, Athens, Greece, June 2011.
- [8] X Zhang, L Chen, M Wang, "Efficient multi-way theta-join processing using MapReduce", In VLDB, Istanbul, Turkey, August 2012.
- [9] V. Borkar, "A Toolkit for the Efficient Processing of Big Data on Large Clusters", In VLDB, Istanbul, Turkey, August 2012.
- [10] B Bahmani, R Kumar and S Vassilvitskii, "Densest subgraph in streaming and mapreduce", In VLDB, Istanbul, Turkey, August 2012.
- [11] Urbani, J., Kotoulas, S., Oren, E. and Van Harmelen, F., "Scalable distributed reasoning using mapreduce", In The Semantic Web-ISWC 2009. Springer Berlin Heidelberg.
- [12] Verma, A., Llorca, X., Goldberg, D. E. and Campbell, R. H., "Scaling genetic algorithms using mapreduce", In ISDA, Pisa, Italy, November 2009.
- [13] Marinelli, E. E. (2009). Hyrax: cloud computing on mobile devices using MapReduce (No. CMU-CS-09-164). CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE.
- [14] A. Dou, V. Kalogeraki, D. Gunopulos, T. Mielikainen and V.H. Tuulos, "Misco: A MapReduce Framework for Mobile Systems", PETRA 2010 Samos, Greece, June 2010.
- [15] N. Fotiou, G. C. Polyzos and D. Trossen. "Illustrating a Publish-Subscribe Internet Architecture", Telecommunication Systems, Springer, February 2011, available on-line in <http://www.springerlink.com/content/t6m0k022042088t5/fulltext.pdf>