



**ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

**ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**Συλλογή και κατηγοριοποίηση δεδομένων κίνησης, μέσω  
κινητών τηλεφώνων, με στόχο την ανίχνευση δραστηριότητας**

**Τραγοπούλου Σπυριδούλα**

Επιβλέποντες: **Βαρλάμης Ηρακλής**, Λέκτορας  
**Τσερπές Κωνσταντίνος**, Λέκτορας

**ΑΘΗΝΑ**

**ΣΕΠΤΕΜΒΡΙΟΣ 2013**



# Περίληψη

Περίληψη.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ:

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ :



# Abstract

Abstract.

SUBJECT AREA:

KEYWORDS:



# Ευχαριστίες

Ευχαριστίες.

# Περιεχόμενα

<b>Πρόλογος</b>	<b>13</b>
<b>1 Εισαγωγή</b>	<b>15</b>
1.1 Σκοπός της πτυχιακής εργασίας . . . . .	15
1.2 Smartphones και Λειτουργικό Σύστημα Android . . . . .	16
1.2.1 Smartphones . . . . .	16
1.2.2 Λειτουργικό Σύστημα Android . . . . .	18
1.2.2.1 Εκδόσεις Android . . . . .	18
1.2.2.2 Αρχιτεκτονική του Android . . . . .	20
1.2.2.3 Χαρακτηριστικά του Android . . . . .	21
1.3 Εξόρυξη Δεδομένων και Εξόρυξη Δεδομένων σε Κινητές Συσκευές . . . . .	23
1.3.1 Εξόρυξη Δεδομένων (Data Mining) . . . . .	23
1.3.1.1 Μέθοδοι Εξόρυξης Δεδομένων . . . . .	25
1.3.2 Εξόρυξη Δεδομένων σε Κινητές Συσκευές (Mobile Data Mining) . . . . .	26
<b>2 Θεωρητικό Υπόβαθρο</b>	<b>29</b>
2.1 Εφαρμογές Δεδομένων Κίνησης . . . . .	29
2.2 Δεδομένα Κίνησης και Χαρακτηριστικά . . . . .	32
2.2.1 Βασικές Αρχές Δεδομένων Κίνησης . . . . .	32
2.2.2 Χαρακτηριστικά Κίνησης . . . . .	34



2.3	Τεχνικές Αναγνώρισης Κίνησης . . . . .	38
2.3.1	Κατηγοριοποίηση . . . . .	38
2.3.2	Συσταδοποίηση τροχιών κίνησης . . . . .	39
2.3.3	Conditional Random Fields (CRF) . . . . .	40
<b>3</b>	<b>Σχεδιασμός</b>	<b>43</b>
3.1	Το περιβάλλον Eclipse . . . . .	43
3.2	Το ADT plug-in . . . . .	44
3.3	Το εργαλείο Android SDK . . . . .	45
3.3.1	SDK Manager . . . . .	45
3.3.2	Προσομοιωτής κινητής συσκευής (Android Emulator) . . . . .	46
3.3.3	AVD Manager . . . . .	47
3.3.4	DDMS (Dalvik Debug Monitor Server) . . . . .	48
3.3.5	ADB (Android Debug Bridge) . . . . .	48
3.4	Ανάπτυξη Android Εφαρμογής στο περιβάλλον Eclipse . . . . .	49
3.5	Το περιβάλλον Weka . . . . .	49
3.5.1	Χρήση του Weka με κώδικα Java . . . . .	49
3.6	Οι χάρτες Google σε περιβάλλον Android . . . . .	49
<b>4</b>	<b>Υλοποίηση</b>	<b>51</b>
<b>5</b>	<b>Αποτελέσματα</b>	<b>53</b>
<b>6</b>	<b>Συμπεράσματα</b>	<b>55</b>
	<b>Βιβλιογραφία</b>	<b>59</b>

# Κατάλογος Σχημάτων

1.1	Οι εκδόσεις Android που έχουν κυκλοφορήσει. . . . .	19
1.2	Η κατανομή χρήσης των εκδόσεων Android. . . . .	20
1.3	Η αρχιτεκτονική του λειτουργικού συστήματος Android . . . . .	21
1.4	Η διαδικασία ανακάλυψης γνώσης από μια βάση δεδομένων (KDD) . . . . .	24
2.1	Στιγμιότυπα χρήσης της εφαρμογής MyTracks . . . . .	30
2.2	Στιγμιότυπα χρήσης της εφαρμογής Nike+ Running . . . . .	31
2.3	Στιγμιότυπα χρήσης της εφαρμογής RunKeeper . . . . .	32
2.4	Απεικόνιση του space-time cube[8] . . . . .	33
3.1	Η αρχιτεκτονική της πλατφόρμας Eclipse . . . . .	44
3.2	Στιγμιότυπο του SDK Manager. . . . .	45
3.3	Στιγμιότυπο του Android Emulator . . . . .	46
3.4	Στιγμιότυπο του AVD Manager. . . . .	47
3.5	Στιγμιότυπο του DDMS. . . . .	48

## **Κατάλογος Πινάκων**



# Πρόλογος

Πρόλογος.



# Κεφάλαιο 1

## Εισαγωγή

Η τεχνολογία των ασύρματων επικοινωνιών και του ubiquitous computing κατακλύζουν την καθημερινότητα της σύγχρονης κοινωνίας. Τα ασύρματα δίκτυα καλύπτουν την μεγαλύτερη έκταση των αστικών κέντρων, ενώ ο αριθμός των χρηστών έξυπνων τηλεφώνων αυξάνεται καθημερινά. Τα smartphones παρέχουν εφαρμογές καταγραφής διάφορων μορφών δεδομένων κάνοντας χρήση ενσωματωμένων αισθητήρων, όπως Bluetooth, GPS, πυξίδα, φωτογραφική μηχανή και μικρόφωνο. Αυτοί οι αισθητήρες, κάνοντας χρήση ασύρματων δικτύων μπορούν να ανιχνεύσουν την θέση και την κίνηση του χρήστη κατά τη διάρκεια χρήσης του κινητού τηλεφώνου. Κατά συνέπεια, η συλλογή δεδομένων κίνησης από την καθημερινότητα του χρήστη αποτελεί μια απλή διαδικασία, η οποία έχει ως στόχο με την ανάλυση των δεδομένων να διευκολύνει τις διάφορες δραστηριότητες του, παρακολουθώντας τον τρόπο ζωής του. Επιπλέον, η παρακολούθηση και ο έλεγχος της καθημερινής δραστηριότητας του ατόμου βρίσκουν εφαρμογή στην υγειονομική περίθαλψη, καθώς με αυτόν τον τρόπο μπορούν να διαγνωστούν σοβαρά προβλήματα υγείας ή ακόμα να εποπτευθούν ασθενείς που χρειάζονται μετεγχειρητική παρακολούθηση.

### 1.1 Σκοπός της πτυχιακής εργασίας

Οι χρήστες έξυπνων κινητών τηλεφώνων χρησιμοποιούν στην καθημερινότητα τους, εφαρμογές που καταγράφουν δεδομένα (θέσης, δραστηριότητας, χρήσης), τα οποία μπορούν να χρησιμοποιηθούν εφόσον αναλυθούν. Σκοπός της παρούσας πτυχιακής εργασίας είναι η ανάπτυξη μιας εφαρμογής για έξυπνα τηλέφωνα, η οποία θα καταγράφει δεδομένα κίνησης, θα τα αναλύει και θα τα συνδυάζει με δεδομένα συγκοινωνιακού δικτύου με στόχο την ανίχνευση δραστηριότητας.

Η εφαρμογή λειτουργεί σε κινητά με λειτουργικό σύστημα Android, καταγράφει διαρκώς δεδομένα θέσης του χρήστη με χρήση GPS και τα επεξεργάζεται τοπικά κάνοντας χρήση τεχνικών εξόρυξης δεδομένων, προσδιορίζοντας τον τύπο κίνησης του χρήστη μέσα στη μέρα. Πιο συγκεκριμένα, η αναγνώριση της κίνησης υλοποιήθηκε με τη μέθοδο της κατηγοριοποίησης (classification), κάνοντας χρήση των αλγορίθμων κατηγοριοποίησης που παρέχει το λογισμικό Weka, όπως ο C4.5 (j48) και ο Random Forest. Για την εκπαίδευση του συστήματος, χρησιμοποιήθηκαν δεδομένα εκπαίδευσης καταγεγραμμένα από την ίδια εφαρμογή.

## **1.2 Smartphones και Λειτουργικό Σύστημα Android**

### **1.2.1 Smartphones**

Με τον όρο έξυπνο τηλέφωνο (smartphone) αναφερόμαστε σε ένα κινητό τηλέφωνο βασισμένο σε ένα λειτουργικό σύστημα κινητής τηλεφωνίας με περισσότερη προηγμένη υπολογιστική ικανότητα και συνδεσιμότητα σε σχέση με ένα απλό κινητό τηλέφωνο. Τα σύγχρονα smartphones αποτελούν πολυχρηστικές συσκευές καθώς περιλαμβάνουν λειτουργίες media players, ψηφιακές φωτογραφικές μηχανές, πλοήγηση GPS, οθόνες αφής υψηλής ανάλυσης και web browsers που εμφανίζουν τυποποιημένες ιστοσελίδες, καθώς και βελτιστοποιημένες ιστοσελίδες για κινητά. Η πρόσβαση σε δεδομένα υψηλής ταχύτητας παρέχεται μέσω Wi-Fi και μέσω ασύρματων ευρυζωνικών υπηρεσιών. Τα λειτουργικά συστήματα (OS) των κινητών τηλεφώνων που χρησιμοποιούνται από τα σύγχρονα σμαρττηφονες περιλαμβάνουν το Android της Google, το iOS της Apple, το Symbian της Nokia, το BlackBerry OS της RIM, το Firefox OS της Mozilla και το Ubuntu Phone της Canonical Ltd's.[25]

### **Αισθητήρες Smartphones**

#### **Global Positioning System (GPS)**

Το GPS (Global Positioning System), Παγκόσμιο Σύστημα Θεσιθεσίας είναι ένα παγκόσμιο σύστημα εντοπισμού θέσης, το οποίο βασίζεται σε εικοσιτέσσερις δορυφόρους της Γης, οι οποίοι διαθέτουν ειδικές συσκευές ονομάζονται δέκτες GPS. Οι δέκτες αυτοί παρέχουν ακριβείς πληροφορίες για τη θέση ενός σημείου, το υψόμετρό του, την ταχύτητα και την κατεύθυνση της κίνησης του. Το GPS παρέχει πληροφορίες τοποθεσίας και χρόνου σε όλες τις καιρικές συνθήκες, οπουδήποτε πάνω ή κοντά στη Γη, όπου υπάρχει ανεμπόδιστη οπτική επαφή με τέσσερις ή περισσότερους δορυφόρους GPS. Συντηρείται από την κυβέρνηση των Ηνωμένων Πολιτειών και



είναι ελεύθερα προσβάσιμο με ένα δέκτη GPS. Το GPS δημιουργήθηκε αναπτύχθηκε το 1973 από το Υπουργείο Άμυνας των ΗΠΑ με σκοπό να ξεπεράσει τους περιορισμούς των προηγούμενων συστημάτων πλοήγησης. Αρχικά ονομάστηκε "NAVSTAR GPS" (Navigation Signal Timing and Ranging Global Positioning System) και τέθηκε σε πλήρη λειτουργία το 1994. [22] Για τον προσδιορισμό της θέσης του χρήστη με γεωγραφικό μήκος (longitude) και γεωγραφικό πλάτος (latitude) απαιτείται σύνδεση σε τρεις δορυφόρους, ενώ για τον προσδιορισμό του υψόμετρου απαιτούνται τέσσερις δορυφόροι.

Η πλειοψηφία των δεκτών GPS βρίσκονται σε κινητά τηλέφωνα, με ποικίλους βαθμούς κάλυψης και προσβασιμότητας των χρηστών. Στα περισσότερα smartphones υπάρχει διαθέσιμο λογισμικό πλοήγησης, καθώς και σε ορισμένα τηλέφωνα που διαθέτουν Java που τους επιτρέπει να χρησιμοποιούν εσωτερικό ή εξωτερικό δέκτη GPS. Ακόμα, μερικά κινητά τηλέφωνα χρησιμοποιούν A-GPS (assisted GPS), το οποίο όμως υπολειτουργεί όταν βρίσκεται εκτός του δικτύου κινητής τηλεφωνίας. Παράλληλα, μερικά άλλα χρησιμοποιούν ένα υβριδικό σύστημα εντοπισμού θέσης που μπορούν να χρησιμοποιήσουν άλλα σήματα, όταν τα σήματα GPS είναι ανεπαρκή.[23]

Στα κινητά τηλέφωνα, η θέση μπορεί να οριστεί είτε με χρήση GPS, είτε με τον τριγωνισμό της απόστασης από την κεραία της κινητής τηλεφωνίας ή το ασύρματο δίκτυο (Wi-Fi) που χρησιμοποιεί ο χρήστης, είτε θεωρώντας ως τοποθεσία του χρήστη την κεραία της κινητής τηλεφωνίας ή του ασύρματου δικτύου. Τα smartphones έχουν τη δυνατότητα να επιλέγουν κάθε φορά τον βέλτιστο τρόπο για τον ορισμό της θέσης του χρήστη.[2]

### **Επιταχυνσιόμετρο (accelerometer)**

Το επιταχυνσιόμετρο είναι μια συσκευή που μετρά την επιτάχυνση. Η επιτάχυνση που μετράει δεν είναι απαραίτητως η επιτάχυνση με βάση τις συντεταγμένες (ρυθμός μεταβολής της ταχύτητας). Αντίθετα, το επιταχυνσιόμετρο υπολογίζει την επιτάχυνση που συνδέεται με το φαινόμενο του βάρους οποιασδήποτε μάζας σε αδράνεια σε σχέση με τη διάταξη της συσκευής.

Μερικά smartphones, περιέχουν επιταχυνσιόμετρα για τον έλεγχο της διεπαφής χρήστη. Συχνά το επιταχυνσιόμετρο χρησιμοποιείται για να παρουσιάσει θέα στο γύρω τοπίο ή πορτρέτο της οθόνης της συσκευής, με βάση την διάταξη της συσκευής. Η 5η και 6η γενιά Apple iPod Nano διαθέτει ένα ενσωματωμένο επιταχυνσιόμετρο και μια εφαρμογή που ονομάζεται Fitness και μπορεί να χρησιμοποιηθεί για να καταγράψει τα βήματα κατά το περπάτημα ή το τρέξιμο.[20]

### **Μαγνητόμετρο (magnetometer)**

Το μαγνητόμετρο είναι ένα όργανο μέτρησης της δύναμης και σε ορισμένες περιπτώσεις, της κατεύθυνσης των μαγνητικών πεδίων. Πολλά smartphones διαθέτουν μαγνητόμετρα και παρέχουν εφαρμογές που λειτουργούν ως πυξίδες. Άλλες συσκευές τηλεφώνων, χρησιμοποιούν μαγνητόμετρα τριών αξόνων, τα οποία δεν είναι ευαίσθητα στον προσανατολισμό ή την ανύψωση της συσκευής. Ερευνητές στην Deutsche Telekom έχουν χρησιμοποιήσει μαγνητόμετρα ενσωματωμένα σε κινητές συσκευές για να επιτρέπουν 3D αλληλεπίδραση του χρήστη χωρίς αφή. Η διεπαφή που δημιούργησαν, ονομάζεται MagiTact και παρακολουθεί τις αλλαγές στο μαγνητικό πεδίο γύρω από ένα κινητό τηλέφωνο για να εντοπίσει τις διάφορες χειρονομίες από ένα χέρι που κρατάει ή φοράει ένα μαγνήτη.[24]

## **1.2.2 Λειτουργικό Σύστημα Android**

Το Android είναι λειτουργικό σύστημα το οποίο τρέχει τον πυρήνα του λειτουργικού Linux και έχει σχεδιαστεί για κινητές συσκευές με οθόνη αφής όπως smartphones και tablets. Αρχικά αναπτύχθηκε από την Google και αργότερα από την Open Handset Alliance. Το πρώτο κινητό τηλέφωνο με Android κυκλοφόρησε τον Οκτώβριο του 2008. Η Google δημοσίευσε το μεγαλύτερο μέρος του κώδικα του Android υπό τους όρους της Apache License, μιας ελεύθερης άδειας λογισμικού. Αυτό σημαίνει ότι το Android μπορεί να τροποποιηθεί και να διανεμηθεί ελεύθερα από τους κατασκευαστές συσκευών και τους προγραμματιστές. Επιπλέον, το Android διαθέτει μια μεγάλη κοινότητα προγραμματιστών που αναπτύσσουν εφαρμογές οι οποίες επεκτείνουν τη λειτουργικότητα των συσκευών, γραμμένες σε μια προσαρμοσμένη έκδοση της γλώσσας προγραμματισμού Java για Android.

Αυτοί οι παράγοντες συνέβαλαν στο να γίνει το Android το πιο ευρέως χρησιμοποιούμενο λειτουργικό σύστημα για smartphone και το λογισμικό που επιλέγουν οι περισσότερες εταιρείες λογισμικού για συσκευές υψηλής τεχνολογίας, καθώς έχει χαμηλό κόστος, είναι προσαρμόσιμο και ελαφρύ και δεν απαιτείται ανάπτυξη λογισμικού από το μηδέν. Αυτό έχει ως αποτέλεσμα, παρά το γεγονός ότι αρχικά σχεδιάστηκε για smartphones και tablets, να έχουν δημιουργηθεί εφαρμογές Android για τηλεοράσεις, κονσόλες παιχνιδιών και ψηφιακές κάμερες.[21]

### **1.2.2.1 Εκδόσεις Android**

Η πρώτη beta έκδοση Android κυκλοφόρησε στις 5 Νοεμβρίου 2007. Από τότε έχουν κυκλοφορήσει πολλές εκδόσεις και έχουν γίνει ενημερώσεις οι οποίες περιγράφονται σε κάθε έκδοση

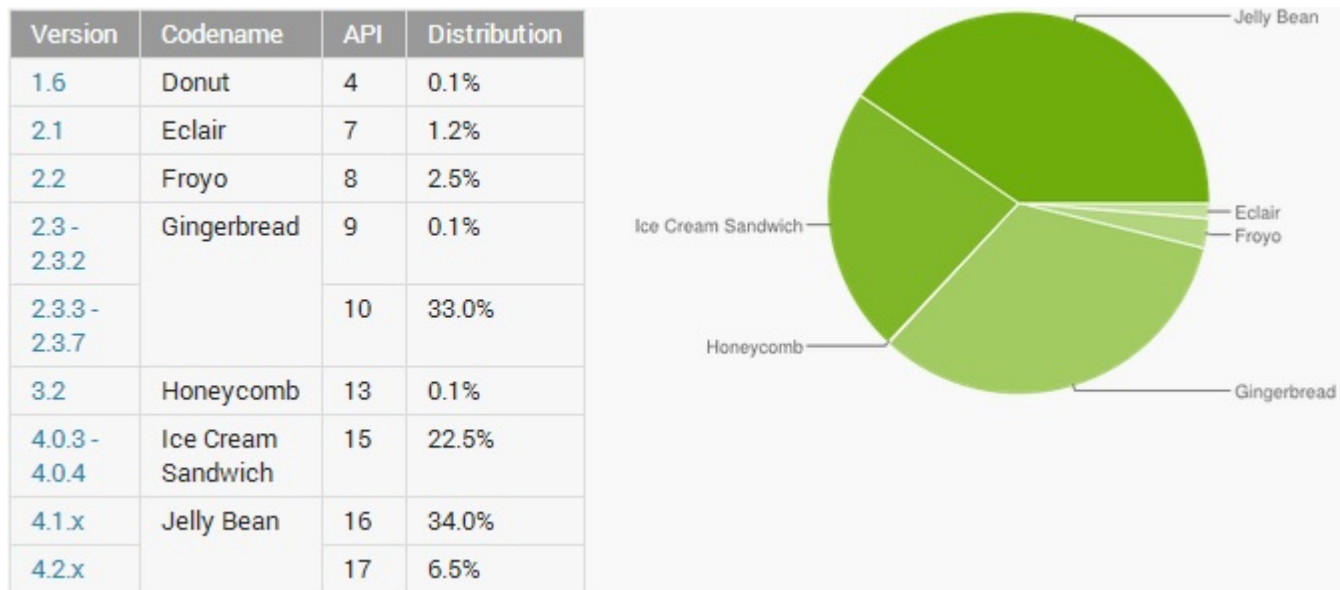
με έναν αριθμό που ονομάζεται API Level και προσδιορίζει το API framework που υποστηρίζεται. Στον πίνακα 1.1 παρουσιάζονται όλες οι εκδόσεις Android που έχουν κυκλοφορήσει, με το API Level που υποστηρίζουν[19].

Platform Version	API Level	VERSION_CODE
Android 4.3	18	JELLY_BEAN_MR2
Android 4.2, 4.2.2	17	JELLY_BEAN_MR1
Android 4.1, 4.1.1	16	JELLY_BEAN
Android 4.0.3, 4.0.4	15	ICE_CREAM_SANDWICH_MR1
Android 4.0, 4.0.1, 4.0.2	14	ICE_CREAM_SANDWICH
Android 3.2	13	HONEYCOMB_MR2
Android 3.1.x	12	HONEYCOMB_MR1
Android 3.0.x	11	HONEYCOMB
Android 2.3.4 Android 2.3.3	10	GINGERBREAD_MR1
Android 2.3.2 Android 2.3.1 Android 2.3	9	GINGERBREAD
Android 2.2.x	8	FROYO
Android 2.1.x	7	ECLAIR_MR1
Android 2.0.1	6	ECLAIR_0_1
Android 2.0	5	ECLAIR
Android 1.6	4	DONUT
Android 1.5	3	CUPCAKE
Android 1.1	2	BASE_1_1
Android 1.0	1	BASE

Σχήμα 1.1: Οι εκδόσεις Android που έχουν κυκλοφορήσει.

Τα στατιστικά στοιχεία χρήσης των εκδόσεων Android, τα οποία έχουν καταγραφεί μέχρι και την 1η Αυγούστου 2013, παρουσιάζονται παρακάτω. Οι εκδόσεις που κατέχουν ποσοστό χρήσης μικρότερο από 0,1% δεν φαίνονται στο σχήμα.

Συλλογή και κατηγοριοποίηση δεδομένων κίνησης, μέσω κινητών τηλεφώνων, με στόχο την ανίχνευση δραστηριότητας



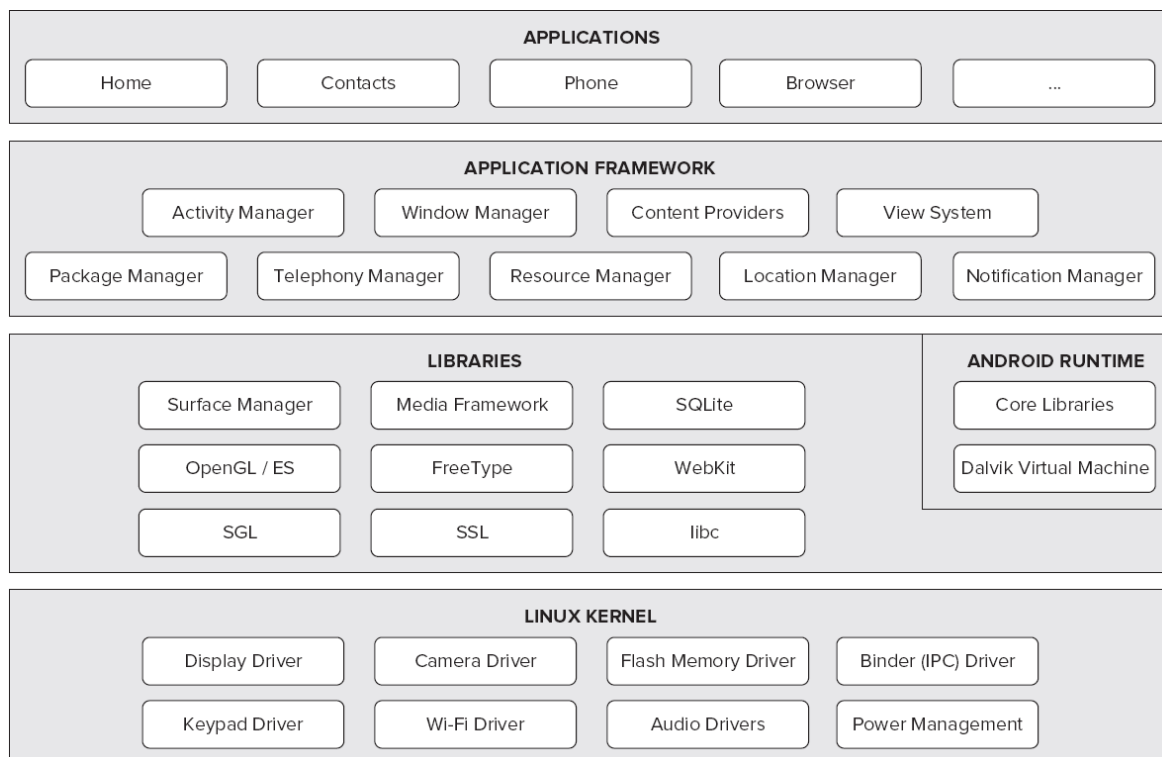
Σχήμα 1.2: Η κατανομή χρήσης των εκδόσεων Android.

#### 1.2.2.2 Αρχιτεκτονική του Android

Το λειτουργικό σύστημα Android αποτελείται από τέσσερα βασικά επίπεδα, τα οποία χωρίζονται σε πέντε τμήματα, όπως φαίνεται στο Σχήμα 1.3[11].

- *Πυρήνας Linux* - Αποτελεί τον πυρήνα στον οποίο είναι βασισμένο το Android. Αυτό το επίπεδο, περιλαμβάνει όλους τους χαμηλού επιπέδου οδηγούς συσκευών για όλα τα hardware εξαρτήματα της συσκευής Android.
- *Βιβλιοθήκες (Libraries)* - Περιλαμβάνει τον κώδικα ο οποίος παρέχει τα βασικά χαρακτηριστικά του Android. Για παράδειγμα, η βιβλιοθήκη WebKit παρέχει λειτουργικότητα για περιήγηση στον ιστό.
- *Android Runtime* - Στο ίδιο επίπεδο με τις βιβλιοθήκες, το Android Runtime διαθέτει βιβλιοθήκες πυρήνα που επιτρέπουν στους προγραμματιστές να αναπτύσσουν εφαρμογές Android κάνοντας χρήση της γλώσσας προγραμματισμού Java. Ακόμα, διαθέτει στην εικονική μηχανή Dalvik, η οποία επιτρέπει σε κάθε εφαρμογή να τρέχει σε ξεχωριστή διεργασία. Η Δαλίκ είναι μια εικονική μηχανή σχεδιασμένη για Ανδرويد και βελτιστοποιημένη για να λειτουργεί σε κινητές συσκευές με μπαταρία και και περιορισμένη μνήμη και επεξεργαστική ισχύ.

- *Framework Εφαρμογών* - Προσφέρει τις ποικίλες δυνατότητες του λειτουργικού Android στους προγραμματιστές εφαρμογών, ώστε να μπορούν να τις χρησιμοποιήσουν στις εφαρμογές που δημιουργούν.
- *Εφαρμογές* - Αποτελεί το υψηλότερο επίπεδο του λειτουργικού και διαθέτει όλες τις εφαρμογές που τρέχουν στη συσκευή όπως το τηλέφωνο, την διαχείριση επαφών και τον περιηγητή ιστού.



Σχήμα 1.3: Η αρχιτεκτονική του λειτουργικού συστήματος Android

### 1.2.2.3 Χαρακτηριστικά του Android

Λόγω του ότι το Android αποτελεί λειτουργικό σύστημα ανοιχτού κώδικα και είναι διαθέσιμο για οποιαδήποτε τροποποίηση, δεν περιλαμβάνει καθορισμένες ρυθμίσεις υλικού και λογισμικού. Παρόλα αυτά, υποστηρίζει από μόνο του τις λειτουργίες που περιγράφονται παρακάτω[11].

- *Αποθήκευση Δεδομένων* - Χρήση της SQLite , μιας ελαφριάς βάσης δεδομένων για τις ανάγκες αποθήκευσης.

- *Συνδεσιμότητα* - Υποστηρίζει τεχνολογίες συνδεσιμότητας GSM/EDGE, CDMA, IDEN, EV-DO, UMTS, Bluetooth, LTE, WiMAX και Wi-Fi.
- *Υπηρεσίες Μηνυμάτων* - Υποστηρίζει αποστολή και λήψη SMS και MMS.
- *Περιήγηση στον Ιστό* - Για την περιήγηση στον ιστό διαθέτει φυλλομετρητή βασισμένο στην ανοιχτή τεχνολογία WebKit.
- *Υποστήριξη Πολυμέσων* - Παρέχει υποστήριξη για τις ακόλουθες μορφές πολυμέσων: H.263, H.264 (σε 3GP ή MP4container), MPEG-4 SP, AMR, AMR-WB, AAC, HE-AAC, MP3, MIDI, OGG Vorbis, WAV, JPEG, PNG, GIF, BMP.
- *Υποστήριξη υλικού* - Υποστηρίζει κάμερες στατικής ή κινούμενης εικόνας, οθόνες αφής, GPS, αισθητήρες επιτάχυνσης, μαγνητόμετρα, καθώς και προξυμιτηρ σενσορς.
- *Multitasking* - Υποστηρίζει multitasking εφαρμογές.

Οι περισσότερες λειτουργίες του τηλεφώνου τρέχουν σαν εφαρμογές πάνω στο ενδιάμεσο λογισμικό (middleware) που διαθέτει. Το ενδιάμεσο λογισμικό είναι γραμμένο σε Java και C/C++. Οι εφαρμογές που τρέχουν σε Android είναι γραμμένες σε Java και μεταγλωττίζονται σε προσαρμοσμένο πηγαίο κώδικα που ονομάζεται Dalvik EXecutable (DEX) και εκτελούνται στην εικονική μηχανή Dalvik VM, η οποία είναι σχεδιασμένη για χρήση σε φορητές συσκευές. Οι εφαρμογές επικοινωνούν με τον μηχανισμό binder IPC, ο οποίος παρέχει διαφανή ανταλλαγή μηνυμάτων με χρήση δεμάτων.[6]

Οι εφαρμογές Android αναπτύσσονται στη γλώσσα Java χρησιμοποιώντας το Android Software Development Kit (SDK). Το SDK περιλαμβάνει ένα ολοκληρωμένο σύνολο εργαλείων ανάπτυξης, όπως πρόγραμμα εντοπισμού σφαλμάτων (debugger), βιβλιοθήκες λογισμικού, emulator κινητού τηλεφώνου που βασίζεται σε QEMU, documentation, δείγματα κώδικα και tutorials. Το επίσημο περιβάλλον ανάπτυξης για εφαρμογές Android είναι το Eclipse (IDE) κάνοντας χρήση του Android Development Tools (ADT) plugin. [21]

Ακόμα, το περιβάλλον προγραμματισμού Eclipse επιτρέπει την ενσωμάτωση εργαλείων και τεχνικών εξόρυξης δεδομένων χρησιμοποιώντας το API του Weka. Το Weka προσφέρει ένα πακέτο κλάσεων για προγραμματιστές που επιτρέπει την διαχείριση και την προεπεξεργασία δεδομένων (preprocessing), αλγορίθμους κατηγοριοποίησης (classification) και συσταδοποίησης (clustering) και την αξιολόγηση των επιδόσεων τους (evaluation). Έτσι, συμπεριλαμβάνοντας το πακέτο των κλάσεων του Weka και μέσω κώδικα Java είναι δυνατή η χρήση τεχνικών εξόρυξης γνώσης σε εφαρμογές Android.[26]

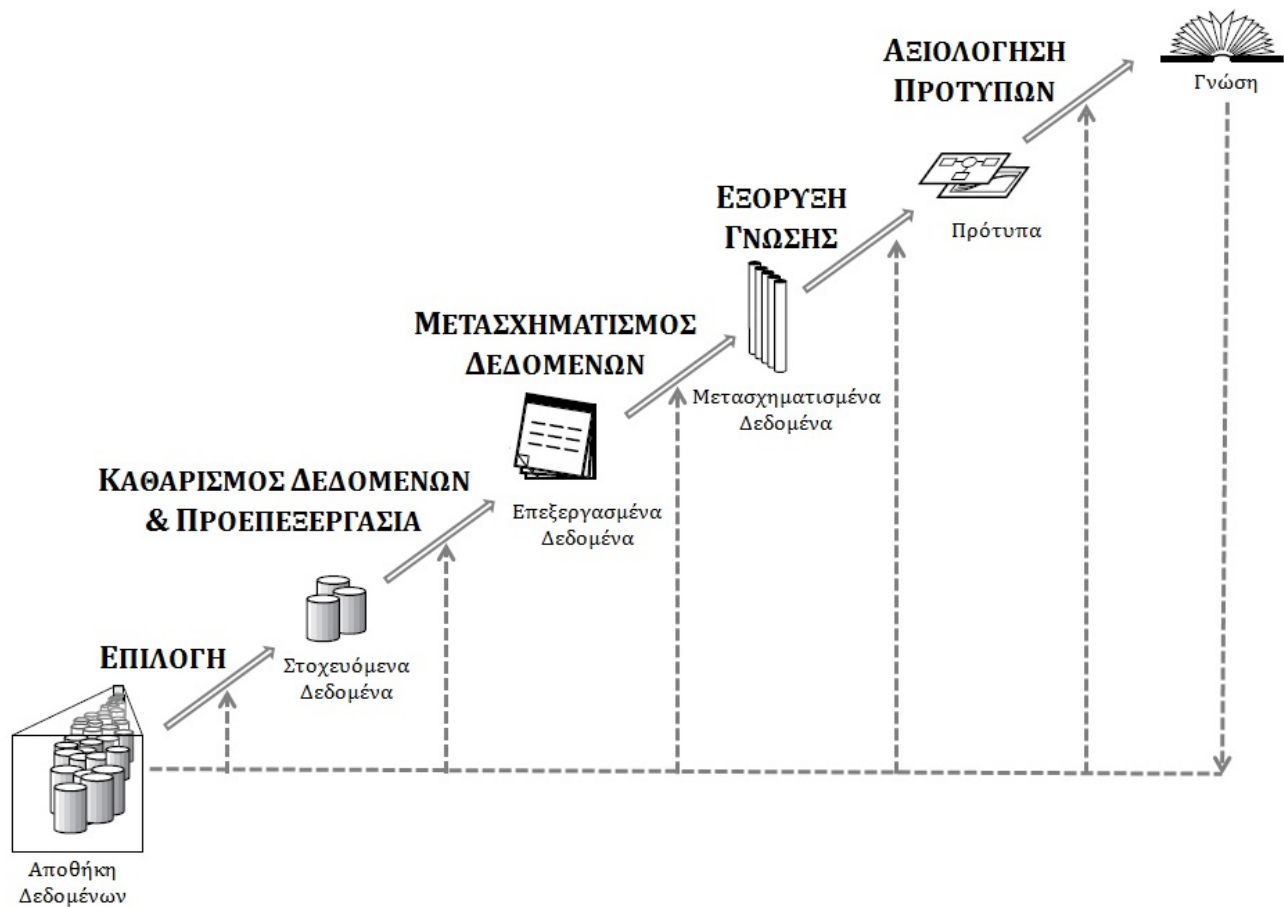
## **1.3 Εξόρυξη Δεδομένων και Εξόρυξη Δεδομένων σε Κινητές Συσκευές**

### **1.3.1 Εξόρυξη Δεδομένων (Data Mining)**

Με τον όρο *εξόρυξη δεδομένων* αναφερόμαστε στη διαδικασία επιλογής, εξερεύνησης και μοντελοποίησης μεγάλου όγκου δεδομένων, με σκοπό την εξαγωγή συμπερασμάτων τα οποία είναι αρχικά άγνωστα, ώστε να προκύψουν σαφή και χρήσιμα αποτελέσματα για τον χρήστη της βάσης δεδομένων[9].

Η ανακάλυψη γνώσης από μια βάση δεδομένων (KDD - Knowledge Discovery from Database), αναφέρεται σε ολόκληρη τη διαδικασία ανακάλυψης χρήσιμης πληροφορίας από μεγάλα σύνολα δεδομένων[1]. Η εξόρυξη δεδομένων αποτελεί το βήμα της KDD διαδικασίας, στο οποίο οι αλγόριθμοι εκμάθησης εφαρμόζονται στα δεδομένα[9].

Η διαδικασία *KDD* είναι μια διαλογική και επαναληπτική διαδικασία που αποτελείται από μια σειρά από τα ακόλουθα βήματα (Σχήμα 1.4) [1].



Σχήμα 1.4: Η διαδικασία ανακάλυψης γνώσης από μια βάση δεδομένων (KDD)

1. *Ανάπτυξη και κατανόηση της περιοχής της εφαρμογής, της υπάρχουσας γνώσης στον τομέα έρευνας και τους τελικούς στόχους.*
2. *Ολοκλήρωση των δεδομένων, συνδυάζοντας πολλαπλές πηγές δεδομένων ώστε να καθοριστεί το σύνολο στο οποίο τελικά η διαδικασία εξόρυξης πρόκειται να εφαρμοστεί.*
3. *Δημιουργία του στόχου - συνόλου δεδομένων, επιλέγοντας το σύνολο δεδομένων (μεταβλητές, δείγματα δεδομένων) στο οποίο η διαδικασία εξόρυξης πρόκειται να εκτελεστεί.*
4. *Καθαρισμός και προεπεξεργασία δεδομένων, εφαρμόζοντας βασικές διαδικασίες όπως η αφαίρεση θορύβου, η συλλογή των απαραίτητων πληροφοριών για τη διαμόρφωση ή τη μέτρηση του θορύβου, η απόφαση σχετικά με τις στρατηγικές διαχείρισης των ελλειπόντων πεδίων δεδομένων.*



5. *Μετασχηματισμός των δεδομένων.* Τα δεδομένα μετασχηματίζονται σε μορφές κατάλληλες για εξόρυξη, κάνοντας χρήση μεθόδων μείωσης διαστάσεων για τη μείωση των υπό εξέταση μεταβλητών ή την εύρεση κατάλληλης αντιπροσώπευσης των δεδομένων χωρίς μεταβλητές.
6. *Επιλογή των στόχων και των αλγορίθμων εξόρυξης δεδομένων.* Σε αυτό το βήμα αποφασίζεται ο στόχος της KDD διαδικασίας, επιλέγοντας τους στόχους εξόρυξης δεδομένων που πρέπει να επιτευχθούν. Ακόμα, επιλέγονται οι μέθοδοι εξόρυξης δεδομένων που θα χρησιμοποιηθούν, λαμβάνοντας υπόψη τις απαιτήσεις και τα γενικά κριτήρια της KDD διαδικασίας.
7. *Εξόρυξη δεδομένων,* εφαρμόζοντας ευφυείς μεθόδους με στόχο την εύρεση προτύπων γνώσης. Τα πρότυπα μπορεί να είναι κανόνες κατηγοριοποίησης, δέντρα, παλινδρόμηση, συσταδοποίηση, κτλ.
8. *Αξιολόγηση των προτύπων.* Τα πρότυπα που προέκυψαν από την εξόρυξη δεδομένων αξιολογούνται με κάποια μέτρα, προκειμένου να προσδιοριστούν τα πρότυπα που αντιπροσωπεύουν καλύτερα την εξαγόμενη γνώση.
9. *Σταθεροποίηση και παρουσίαση της γνώσης,* ενσωματώνοντας τη γνώση στο σύστημα ή απεικονίζοντας την χρησιμοποιώντας τεχνικές αντιπροσώπευσης γνώσης, ώστε να μπορεί να παρουσιαστεί η εξορυγμένη γνώση στον χρήστη.

#### **1.3.1.1 Μέθοδοι Εξόρυξης Δεδομένων**

Η εξόρυξη δεδομένων έχει ως βασικό στόχο την εφαρμογή τεχνικών περιγραφής και πρόβλεψης σε μεγάλα σύνολα δεδομένων. Η *πρόβλεψη* έχει ως στόχο την πρόβλεψη της συμπεριφοράς κάποιων μεταβλητών που παρουσιάζουν ενδιαφέρον και οι οποίες βασίζονται στην συμπεριφορά άλλων μεταβλητών. Η *περιγραφή* στοχεύει στην ανακάλυψη προτύπων και αναπαριστά τα δεδομένα μιας πολύπλοκης βάσης δεδομένων με κατανοητό και αξιοποιήσιμο τρόπο. Όλες οι υπάρχουσες μέθοδοι εξόρυξης δεδομένων έχουν ως βασικό στόχο να προσδιορίσουν και να περιγράψουν τα πρότυπα γνώσης που εξάγονται από ένα σύνολο δεδομένων. Οι κυριότερες μέθοδοι εξόρυξης δεδομένων περιγράφονται παρακάτω.

- *Κατηγοριοποίηση (Classification).* Αποτελεί μια από τις βασικές εργασίες εξόρυξης δεδομένων. Βασίζεται στην εξέταση των χαρακτηριστικών ενός μη-κατηγοριοποιημένου αντικειμένου, το οποίο με βάση αυτά τα χαρακτηριστικά αντιστοιχίζεται σε κάποια από τις

κατηγορίες που έχουν προκαθοριστεί. Η βασική εργασία κατηγοριοποίησης, είναι η δημιουργία ενός μοντέλου το οποίο θα μπορούσε να εφαρμοστεί για να κατηγοριοποιεί μη-κατηγοριοποιημένα δεδομένα. Χρησιμοποιεί ένα σύνολο κατηγοριοποιημένων δεδομένων για την εκπαίδευση του μοντέλου και απαιτεί έναν καλά καθορισμένο ορισμό των κατηγοριών. Οι αλγόριθμοι κατηγοριοποίησης διακρίνονται σε δυο βασικές κατηγορίες τεχνικών, τα Δέντρα Απόφασης (Decision Trees) και τα Νευρωνικά Δίκτυα (Neural Networks).

- *Συσταδοποίηση (Clustering)*. Είναι η διαδικασία καταμερισμού ετερογενών δεδομένων σε ένα σύνολο περισσότερων ετερογενών συστάδων. Σε αντίθεση με την κατηγοριοποίηση, η συσταδοποίηση δεν βασίζεται σε προκαθορισμένες κατηγορίες. Οι εγγραφές των δεδομένων ομαδοποιούνται σε σύνολα με βάση την ομοιότητα που παρουσιάζουν μεταξύ τους.
- *Κανόνες Συσχέτισης (Association Rules)*. Οι κανόνες συσχέτισης στοχεύουν στην ανακάλυψη κρυμμένων συσχετίσεων μεταξύ των γνωρισμάτων ενός συνόλου δεδομένων. Παρέχουν έναν συνοπτικό τρόπο για να εκφραστούν οι ενδεχομένως χρήσιμες πληροφορίες και να γίνουν κατανοητές από τους τελικούς χρήστες.
- *Πρότυπα ακολουθιών (Sequential Patterns)*. Η εξόρυξη πρότυπων ακολουθιών αναφέρεται στην ανίχνευση των συχνά εμφανιζόμενων προτύπων σχετικών με τον χρόνο ή άλλες ακολουθίες. Οι περισσότερες έρευνες στα πρότυπα ακολουθιών επικεντρώνονται σε συμβολικά πρότυπα.
- *Παλινδρόμηση (Regression)*. Αναφέρεται στην εκμάθηση μιας λειτουργίας που εκχωρεί τα δεδομένα σε μια μεταβλητή πρόβλεψης, η οποία παίρνει πραγματικές τιμές.

### 1.3.2 Εξόρυξη Δεδομένων σε Κινητές Συσκευές (Mobile Data Mining)

Ο στόχος της εξόρυξης γνώσης σε κινητές συσκευές (mobile data mining) είναι να παρέχει προηγμένες τεχνικές για την ανάλυση και την παρακολούθηση δεδομένων που καταγράφονται μέσω κινητών συσκευών[18].

Η εξόρυξη δεδομένων σε κινητές συσκευές, έχει να αντιμετωπίσει τα τυπικά ζητήματα της εξόρυξης δεδομένων σε ένα κατανομημένο περιβάλλον ενώ ταυτόχρονα και τους τεχνολογικούς περιορισμούς ενός κινητού τηλεφώνου, όπως δίκτυα με χαμηλό bandwidth, περιορισμένη χωρητικότητα αποθήκευσης, μικρή ισχύς μπαταρίας, αργούς επεξεργαστές και μικρές οθόνες για την απεικόνιση των αποτελεσμάτων[16].

Ο τομέας της εξόρυξης δεδομένων σε κινητές συσκευές περιλαμβάνει διάφορα σενάρια εφαρμογών, στις οποίες ένα κινητό τηλέφωνο μπορεί να καταγράφει δεδομένα, να τα αναλύει, να αποτελεί τον client ενός απομακρυσμένου server που υλοποιεί την εξόρυξη δεδομένων ή να υλοποιεί τον συνδυασμό αυτών των λειτουργιών. Πιο συγκεκριμένα, μπορούμε να διακρίνουμε τρία βασικά σενάρια για την εξόρυξη δεδομένων σε κινητές συσκευές[18].

1. Η κινητή συσκευή χρησιμοποιείται σαν τερματικό για την πρόσβαση σε έναν απομακρυσμένο server ο οποίος παρέχει υπηρεσίες εξόρυξης δεδομένων. Ο server αναλύει τα δεδομένα που είναι αποθηκευμένα σε μια τοπική ή κατανεμημένη βάση δεδομένων και στέλνει τα αποτελέσματα της διαδικασίας εξόρυξης στην συσκευή για να τα απεικονίσει.
2. Τα δεδομένα που καταγράφηκαν μέσω μιας κινητής συσκευής, αποστέλλονται σε έναν απομακρυσμένο server και αποθηκεύονται σε μια τοπική βάση δεδομένων. Τα δεδομένα αναλύονται κάνοντας χρήση συγκεκριμένων αλγορίθμων εξόρυξης δεδομένων και τα αποτελέσματα χρησιμοποιούνται για να ληφθούν αποφάσεις για κάποιο συγκεκριμένο σκοπό.
3. Η ανάλυση των δεδομένων υλοποιείται στις κινητές συσκευές. Βέβαια λόγω της μειωμένης επεξεργαστικής ισχύος και χωρητικότητας των σύγχρονων συσκευών, δεν είναι πάντα εφικτή η υλοποίηση όλων των αλγορίθμων εξόρυξης δεδομένων σε μια κινητή συσκευή. Ωστόσο, κάποια βήματα της διαδικασίας εξόρυξης, όπως η επιλογή και η προεπεξεργασία των δεδομένων μπορούν να εκτελεστούν σε μια τέτοια συσκευή, καθώς επίσης και μερικοί αλγόριθμοι εξόρυξης γνώσης που δεν απαιτούν μεγάλη επεξεργαστική ισχύ.

Μια πιο σύγχρονη προσέγγιση εξόρυξης δεδομένων σε κινητές συσκευές παρουσιάζουν οι Steinbauer και συν.[14], η οποία βασίζεται σε τεχνολογία cloud. Ο στόχος τους είναι να παρέχουν εργαλεία και τεχνικές για την ανάλυση δεδομένων κοινωνικών δικτύων σε σχεδόν πραγματικό χρόνο. Τα δεδομένα παράγονται και καταγράφονται από τους αισθητήρες της κινητής συσκευής και έπειτα αποστέλλονται σε ένα δυναμικό μοντέλο cloud για περαιτέρω επεξεργασία.



## **Κεφάλαιο 2**

### **Θεωρητικό Υπόβαθρο**

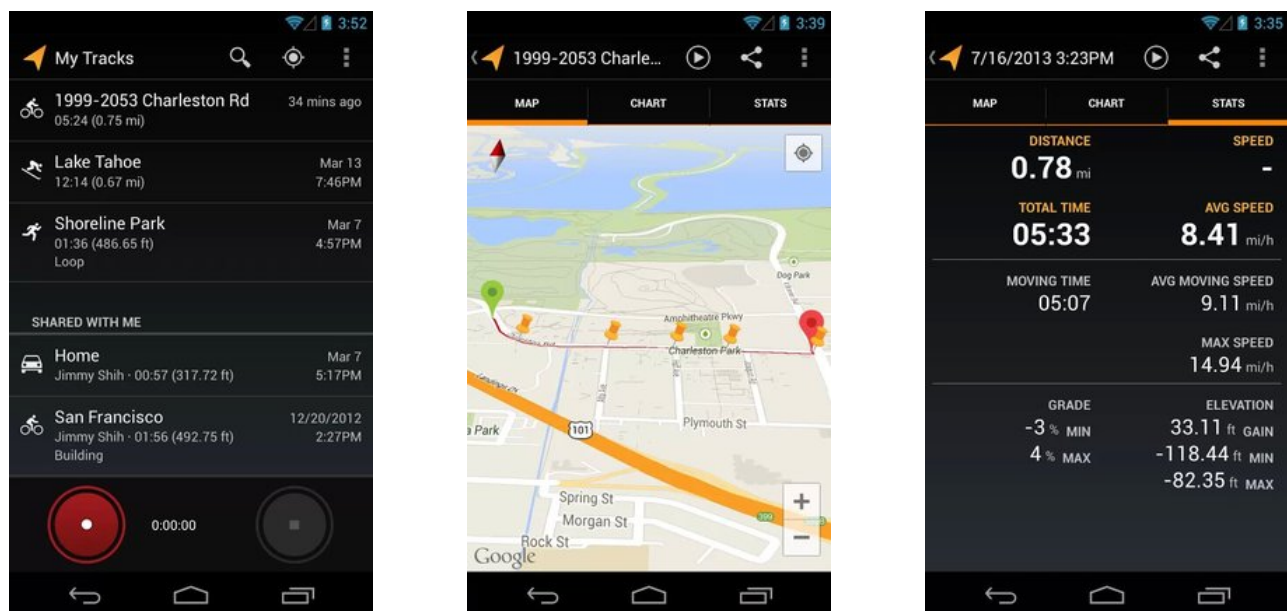
#### **2.1 Εφαρμογές Δεδομένων Κίνησης**

Παρακάτω παρουσιάζονται μερικές υπάρχουσες εφαρμογές για Android οι οποίες χρησιμοποιούν δεδομένα θέσης για να καταγράψουν την δραστηριότητα του χρήστη.

##### **My Tracks**

Η εφαρμογή My Tracks της Google καταγράφει τη διαδρομή, την ταχύτητα, την απόσταση και το υψόμετρο του χρήστη καθώς περπατάει, τρέχει, κάνει ποδήλατο ή οποιαδήποτε άλλη υπαίθρια δραστηριότητα.

Συλλογή και κατηγοριοποίηση δεδομένων κίνησης, μέσω κινητών τηλεφώνων, με στόχο την ανίχνευση δραστηριότητας

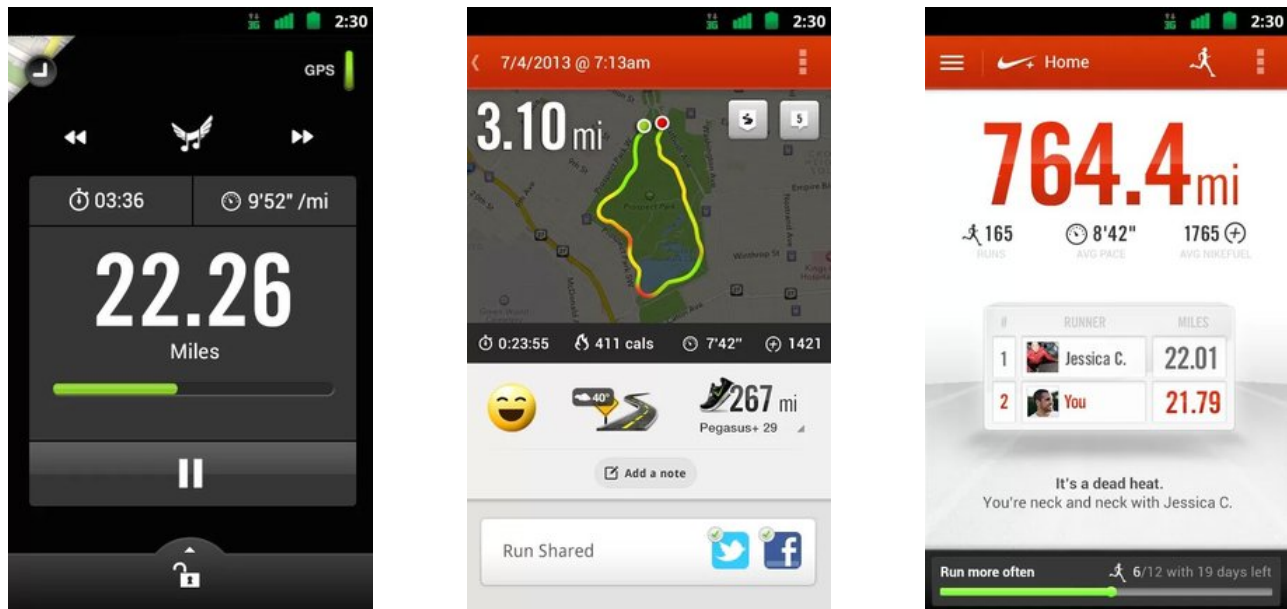


Σχήμα 2.1: Στιγμιότυπα χρήσης της εφαρμογής MyTracks

Ενώ καταγράφει δεδομένα, επιτρέπει στον χρήστη να δει τα δεδομένα κίνησης του, να χαρακτηρίσει την κίνηση που κάνει εκείνη τη στιγμή, ενώ ταυτόχρονα ακούει ανακοινώσεις για την πρόοδο των επιδόσεων του. Επίσης, εμφανίζει γραφήματα για την μεταβολή της ταχύτητας και του υψομέτρου σε προκαθορισμένες διαδρομές. Χρησιμοποιεί GPS για να καταγράψει γεωγραφικά δεδομένα και στατιστικά ταχύτητας αλλά και εξωτερικούς βιομετρικούς αισθητήρες (Zephyr HxM, Polar WearLink, ANT+) για να καταγράψει τον καρδιακό ρυθμό και την ταχύτητα.

## Nike+ Running

Η εφαρμογή Nike+ Running της Nike χρησιμοποιεί το GPS και τον αισθητήρα επιτάχυνσης (accelerometer) για να καταγράψει με ακρίβεια την απόσταση, τον ρυθμό και τον χρόνο της διαδρομής του χρήστη.

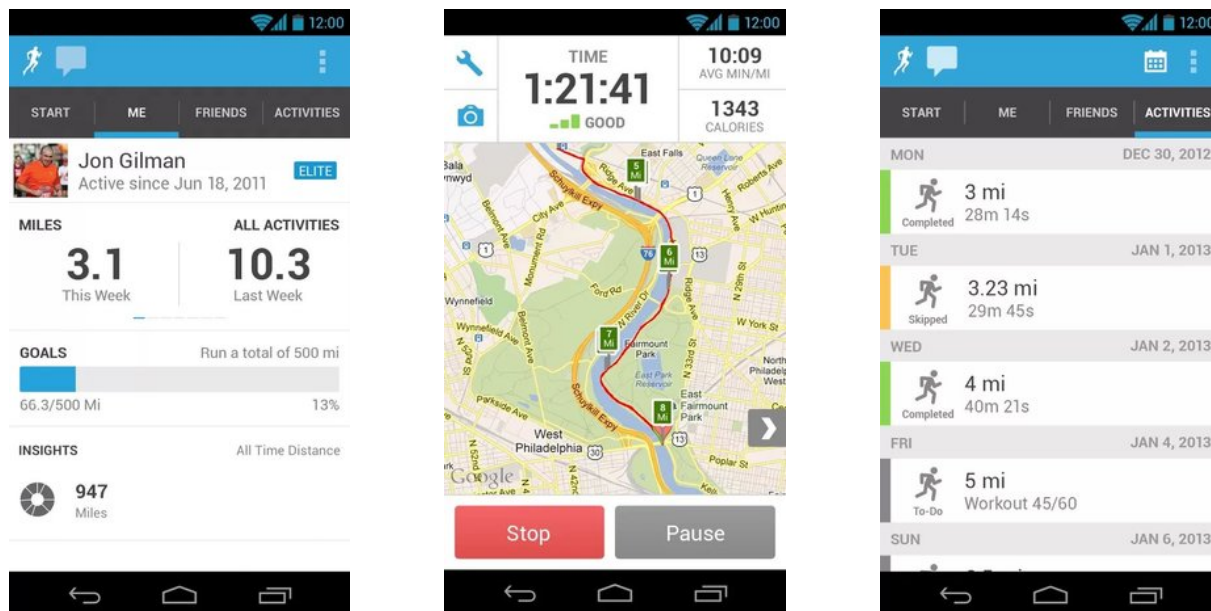


Σχήμα 2.2: Στιγμιότυπα χρήσης της εφαρμογής Nike+ Running

Έχει ως σκοπό να δώσει κίνητρο στον χρήστη για φυσική δραστηριότητα και για αυτό το λόγο του επιτρέπει να προσθέτει φίλους του να παρακολουθεί την δραστηριότητα τους και να μοιράζεται μαζί τους τις διαδρομές που έχει κάνει, ενώ διαθέτει επιλογή ενώ ο χρήστης κινείται να ακούει τα αγαπημένα του τραγούδια.

## RunKeeper

Η εφαρμογή Runkeeper της FitnessKeeper έχει ως στόχο να κάνει ευχάριστη την φυσική δραστηριότητα του χρήστη. Λειτουργεί σε κινητά Android και χρησιμοποιεί το GPS της συσκευής για να καταγράψει τρέξιμο, περπάτημα, ποδηλασία και άλλες δραστηριότητες και εμφανίζει στατιστικά για τον ρυθμό, την απόσταση, τον χρόνο, τις θερμίδες που καίει και τον καρδιακό του ρυθμό. Ο χρήστης μπορεί να ενημερώνεται για τα στατιστικά και την πρόοδο του μέσω των ακουστικών του ενώ ταυτόχρονα ακούει συμβουλές προπόνησης. Επίσης, έχει τη δυνατότητα να ακούει μουσική κατά τη διάρκεια της άσκησης και να τραβάει φωτογραφίες για να τις μοιράζεται. Ακόμα, εφαρμογή κρατάει ιστορικό των δραστηριοτήτων και ενημερώνει τον χρήστη όταν φτάσει σε νέο ρεκόρ επίδοσης, ενώ του προτείνει πλάνα άσκησης για να φτάσει σε συγκεκριμένη φυσική κατάσταση.



Σχήμα 2.3: Στιγμιότυπα χρήσης της εφαρμογής RunKeeper

## 2.2 Δεδομένα Κίνησης και Χαρακτηριστικά

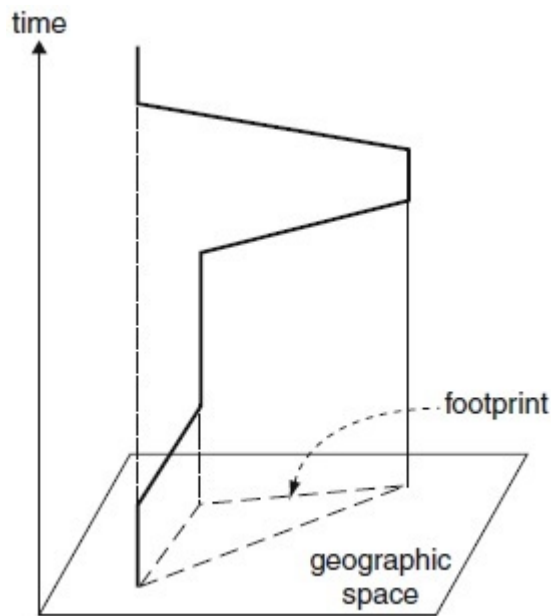
Στις μέρες μας, οι καθημερινές δραστηριότητες των ανθρώπων δημιουργούν συνεχώς ψηφιακά ίχνη μέσω των ασύρματων δικτύων των κινητών τηλεφώνων. Οι τεχνολογίες προσδιορισμού θέσης όπως το GPS, το UMTS και το GSM, συνεχώς βελτιώνονται ως προς την ακρίβεια της θέσης του χρήστη. Αυτό έχει ως αποτέλεσμα, οι ανθρώπινες δραστηριότητες μπορούν να αναγνωριστούν χρησιμοποιώντας τα δεδομένα κίνησης που καταγράφονται μέσω των κινητών τηλεφώνων. Όμως, με ποιο τρόπο μπορεί να καταγραφεί η κίνηση του χρήστη και τι είδους δεδομένα απαιτούνται για τον καλύτερο προσδιορισμό της κίνησης;

### 2.2.1 Βασικές Αρχές Δεδομένων Κίνησης

Μια από τις πρώτες προοπτικές ανάλυσης των προτύπων ανθρώπινης δραστηριότητας και της κίνησης στον χωροχρόνο είναι η χρονο-γεωγραφία (time geography). Αναπτύχθηκε από μια ομάδα Σουηδών γεωγράφων το 1970, με σημαντικότερο εκπρόσωπο τον Torsten Hägerstrand. Η χρονο-γεωγραφία έχει χρησιμοποιηθεί από γενιές κοινωνικών επιστημόνων, ιδιαίτερα γεωγράφων και ερευνητών μεταφορών, για την περιγραφή και την ανάλυση των ανθρώπινων δραστηριοτήτων στο χώρο-χρόνο. Θεωρεί και απεικονίζει τις δραστηριότητες ενός ατόμου σε ένα 24-ωρο ως μία συνεχή χρονική ακολουθία στο γεωγραφικό χώρο. Ο αριθμός και οι τοποθεσίες



των καθημερινών δραστηριοτήτων ενός ατόμου περιορίζονται από τον διαθέσιμο χρόνο και από ποικίλες υποχρεωτικές δραστηριότητες (π.χ. εργασία). Η τροχιά της κίνησης του χρήστη θεωρείται μια διαδρομή στον χωρόχρονο και αναπαρίσταται στον τρισδιάστατο χώρο χρησιμοποιώντας τον οριζόντιο άξονα για την αναπαράσταση του γεωγραφικού χώρου και τον κάθετο άξονα για τον χρόνο. Η αναπαράσταση αυτή ονομάζεται κύβος χωροχρόνου (space-time cube) και απεικονίζεται στο Σχήμα 2.4 [8][15]



Σχήμα 2.4: Απεικόνιση του space-time cube[8]

Η γραμμή αναπαριστά τις κινήσεις μιας οντότητας, για παράδειγμα, ας θεωρήσουμε έναν εργαζόμενο, ο οποίος αρχικά ήταν στο σπίτι, μετά πήγε στη δουλειά του και έμεινε εκεί για ένα χρονικό διάστημα, μετά πήγε στο σουπερμάρκετ για ψώνια και αφού έμεινε κι εκεί για λίγο, γύρισε σπίτι. Οι κάθετες γραμμές ερμηνεύονται σαν την παραμονή του χρήστη σε συγκεκριμένο μέρος (σπίτι, εργασία, σουπερμάρκετ). Τα κεκλιμένα τμήματα της γραμμής υποδεικνύουν κινήσεις. Όσο πιο αργή είναι η κίνηση, τόσο πιο απότομη είναι η γραμμή. Η ευθεία γραμμή δείχνει ότι το άτομο κινείται με σταθερή ταχύτητα, η οποία είναι συνήθως μια προσέγγιση της πραγματικής συμπεριφοράς. Η απεικόνιση αυτής της αναπαράστασης σε χάρτη είναι το αποτύπωμα της διαδρομής που ακολούθησε ο χρήστης.

### 2.2.2 Χαρακτηριστικά Κίνησης

Για την αναγνώριση της δραστηριότητας του χρήστη πρέπει να εξεταστούν πολλοί παράγοντες της κίνησης. Αρχικά, τα χαρακτηριστικά της κάθε κίνησης όπως για παράδειγμα η ταχύτητα και η κατεύθυνση, η τροχιά της κίνησης που ακολουθεί ο χρήστης αλλά και τα χαρακτηριστικά του χρήστη. Παράλληλα, πολλά συμπεράσματα μπορούν να εξαχθούν για την φύση της κίνησης παρατηρώντας το περιβάλλον της και τα συμβάντα που εξελίσσονται σε αυτό. [8]

Η έννοια της *κίνησης* μπορεί να οριστεί σαν την μεταβολή της φυσικής θέσης μιας οντότητας σε σχέση με ένα σύστημα αναφοράς στο οποίο μπορεί να προσδιοριστεί η θέση της οντότητας. Το σύστημα αναφοράς συνήθως είναι ένας γεωγραφικός χώρος.

Η κίνηση μιας οντότητας αποτελείται από *τροχιές*, δηλαδή από τις διαδρομές που ακολουθεί η οντότητα στον χώρο κατά την κίνηση της. Μια διαδρομή δεν δημιουργείται ποτέ στιγμιαία, αντίθετα απαιτεί ένα χρονικό διάστημα. Για αυτό τον λόγο, η τροχιά στην οποία κινείται μια οντότητα πρέπει να συνδέεται με τον χρόνο και μπορεί να περιγραφεί από μια σειρά εγγραφών της μορφής (χρονική στιγμή, θέση). Ο τρόπος καταγραφής μιας τροχιάς ποικίλει, καθώς διαφέρει ανάλογα με τον τρόπο παρατήρησης της κίνησης.

#### Τρόποι καταγραφής κίνησης

- *Χρονική.* Η καταγραφή της θέσης της οντότητας γίνεται ανά τακτά χρονικά διαστήματα (π.χ. ανά λεπτό).
- *Με βάση την μεταβολή θέσης.* Γίνεται καταγραφή όταν η θέση της οντότητας διαφέρει από την προηγούμενη.
- *Με βάση τη θέση κοντά σε συγκεκριμένη τοποθεσία.* Καταγραφή της θέσης όταν η οντότητα πλησιάζει σε μια συγκεκριμένη τοποθεσία.
- *Με βάση κάποιο συμβάν.* Η θέση της οντότητας καταγράφεται όταν συμβεί κάποιο συγκεκριμένο γεγονός, συνήθως αφορά δραστηριότητες των ίδιων οντοτήτων.

Όσον αφορά στα χαρακτηριστικά της κίνησης που μελετούνται, μπορούν να διακριθούν σε δυο κατηγορίες ανάλογα με το αν αναφέρονται σε χρονική στιγμή ή σε χρονικό διάστημα.

#### Χαρακτηριστικά χρονικής στιγμής της κίνησης

- Timestamp της συγκεκριμένης χρονικής στιγμής

- Θέση της οντότητας στον χώρο
- Κατεύθυνση
- Ταχύτητα
- Αλλαγή κατεύθυνσης
- Επιτάχυνση (αλλαγή ταχύτητας)
- Συνολικός χρόνος κίνησης
- Συνολική απόσταση που έχει διανυθεί

### **Χαρακτηριστικά χρονικού διαστήματος της κίνησης**

- Γεωμετρικό σχήμα της κίνησης στον χώρο
- Απόσταση που έχει διανυθεί στον χώρο
- Χρονική διάρκεια της κίνησης
- Διάνυσμα της κίνησης (από την αρχική στην τελική θέση)
- Μέση και μέγιστη τιμή της ταχύτητας
- Δυναμική συμπεριφορά της ταχύτητας (σταθερή ταχύτητα, επιτάχυνση, επιβράδυνση, μη-δενική ταχύτητα)
- Δυναμική συμπεριφορά της κατεύθυνσης (ευθεία, καμπυλόγραμμη, κυκλική κίνηση)

### **Χώρος**

Ως χώρος μπορεί να θεωρηθεί το σύνολο από τοποθεσίες και μέρη, με την ιδιότητα να μπορεί να οριστεί η απόσταση μεταξύ τους. Για τη διάκριση των διαφορετικών τοποθεσιών στον χώρο πρέπει να θεωρηθεί ένα σύστημα αναφοράς, όπως ένα σύστημα συντεταγμένων. Ανάλογα με τις απαιτήσεις του προβλήματος, ο χώρος μπορεί να θεωρηθεί ως δισδιάστατος ή τρισδιάστατος, δηλαδή ένα σημείο μπορεί να περιγραφεί από δύο ή τρεις συντεταγμένες αντίστοιχα.

Ο φυσικός χώρος είναι συνεχής, το οποίο σημαίνει ότι αποτελείται από άπειρο αριθμό από τοποθεσίες. Από την άλλη πλευρά, κάποιες φορές μπορεί να είναι χρήσιμο να θεωρούμε τον

χώρο διακριτό ή από πεπερασμένο αριθμό από τοποθεσίες. Σε κάποιες περιπτώσεις, η διακριτοποίηση του χώρου μπορεί να είναι απαραίτητη, όταν οι θέσεις των οντοτήτων δεν μπορούν να προσδιοριστούν με ακρίβεια αλλά με την έννοια ευρύτερης περιοχής όπως για παράδειγμα περιοχές δικτύου κινητής τηλεφωνίας, περιοχές μιας πόλης ή χώρες.

Ο χώρος μπορεί να δομηθεί με διάφορους τρόπους για να επιτευχθεί καλύτερη καταγραφή της κίνησης μιας οντότητας. Ένας τρόπος είναι η ιεραρχική διαίρεση, για παράδειγμα μια χώρα χωρίζεται σε επαρχίες, οι επαρχίες σε δήμους και οι κοινότητες σε επαρχίες. Περιοχές μπορούν να προκύψουν ακόμα από γεωγραφική διαίρεση, δηλαδή δημιουργώντας κελιά στον χώρο με συγκεκριμένο μέγεθος (π.χ.  $1 \text{ km}^2$ ). Επίσης, ένας πολύ κοινός τρόπος δόμησης του φυσικού χώρου είναι με βάση το οδικό δίκτυο.

Γενικά, οι πιθανοί τρόποι με τους οποίους μπορεί να προσδιοριστεί η θέση μιας οντότητας στον γεωγραφικό χώρο περιγράφονται παρακάτω.

- *Αναφορά με βάση συντεταγμένες.* Οι θέσεις ορίζονται με αριθμούς που αντιστοιχούν στη γραμμική ή γωνιακή απόσταση από καθορισμένους άξονες ή γωνίες.
- *Αναφορά με βάση τη διαίρεση του χώρου.* Αναφέρεται σε περιοχές μιας γεωμετρική ή σημασιολογική διαίρεση του χώρου, ενδεχομένως ιεραρχική.
- *Γραμμική αναφορά.* Αναφέρονται σε σχετικές θέσεις κατά μήκος γραμμικών αντικειμένων όπως δρόμοι, ποτάμια, αγωγοί. Για παράδειγμα, ονόματα δρόμων καθώς και αριθμούς οπτιών ή κωδικούς δρόμων και αποστάσεις από κάποιο από τα άκρα.

## **Χρόνος**

Από μαθηματική άποψη, ο χρόνος είναι μια συνεχής σειρά στοιχείων με γραμμική διάταξη και αποστάσεις μεταξύ των στοιχείων, όπου τα στοιχεία είναι στιγμές ή θέσεις στο χρόνο. Όπως συμβαίνει και με τον χώρο, απαιτείται ένα σύστημα αναφοράς για τον προσδιορισμό των χρονικών στιγμών στα δεδομένα. Στις περισσότερες περιπτώσεις, ως σύστημα αναφοράς χρησιμοποιείται το Γρηγοριανό ημερολόγιο (Gregorian calendar), και η διαίρεση του χρόνου σε ημέρες, μήνες, ώρες, λεπτά και δευτερόλεπτα. Η ώρα της ημέρας μπορεί να προσδιοριστεί με τη ζώνη ώρας της περιοχής στην οποία καταγράφονται τα δεδομένα ή με βάση την ζώνη ώρας του Greenwich (GMT). Υπάρχουν βέβαια περιπτώσεις, στις οποίες τα δεδομένα αναφέρονται σε σχετικές χρονικές στιγμές, για παράδειγμα ο χρόνος που έχει παρέλθει από την έναρξη μιας διαδικασίας ή της παρατήρησης.

Ο πραγματικός χρόνος περιλαμβάνει επίσης χρονικούς κύκλους που προκύπτουν από την καθημερινή και την ετήσια περιστροφή της γης. Αυτοί οι φυσικοί χρονικοί κύκλοι γίνονται αντιληπτοί με τη φυσική ροή του χρόνου, για παράδειγμα οι ημερομηνίες επαναλαμβάνονται κάθε χρόνο και οι ώρες κάθε μέρα. Εκτός από τους φυσικούς χρονικούς κύκλους, υπάρχουν και αυτοί που σχετίζονται με τις δραστηριότητες των ανθρώπων, για παράδειγμα μια δραστηριότητα μπορεί να γίνεται μια φορά τη μέρα ή τη βδομάδα. Για τη ανάλυση των κινήσεων, είναι πολύ σημαντικό να είναι γνωστοί οι χρονικοί κύκλοι που σχετίζονται με την κίνηση. Οι ιδιότητες κάθε χρονικού κύκλου μπορεί να διαφέρουν και αυτές οι διαφορές μπορεί να έχουν σημαντική επίδραση στις κινήσεις. Για παράδειγμα, οι κινήσεις των ανθρώπων τις καθημερινές διαφέρουν από τα Σαββατοκύριακα.

Όμως, η ετερογένεια των ιδιοτήτων κάθε χρονικής στιγμής δεν μπορεί να εκφραστεί σαφώς στα δεδομένα και για αυτό τον λόγο δεν είναι δυνατό να ληφθούν υπόψη αυτόματα στην ανάλυση των δεδομένων. Η ανάλυση των ιδιοτήτων του χώρου εξαρτάται σημαντικά από την ικανότητα του αναλυτή να χρησιμοποιήσει τις γνώσεις του και είναι απαραίτητο οι μέθοδοι και τα εργαλεία που θα χρησιμοποιήσει για την ανάλυση τους να του δίνουν αυτή τη δυνατότητα.

### **Κινούμενες οντότητες και τα χαρακτηριστικά τους**

Εκτός από τα χαρακτηριστικά της κίνησης που αναφέρθηκαν, οι οντότητες που κινούνται έχουν τα δικά τους χαρακτηριστικά, τα οποία επηρεάζουν την κίνηση τους και πρέπει να ληφθούν υπόψη κατά την ανάλυση των δεδομένων κίνησης. Έτσι, οι κινήσεις των ανθρώπων μπορεί να επηρεάζονται σημαντικά από το επάγγελμα, την ηλικία, την κατάσταση υγείας, την οικογενειακή κατάσταση και από άλλα χαρακτηριστικά. Ακόμα, μπορεί να αναλυθεί ο σκοπός της κίνησης ενός ανθρώπου, καθώς μπορεί έτσι να καθοριστεί η διαδρομή που θα ακολουθήσει και η ταχύτητα κίνησης του. Τα χαρακτηριστικά της κίνησης μπορεί επίσης να εξαρτώνται από τις δραστηριότητες των ανθρώπων ενώ κινούνται. Για παράδειγμα, η κίνηση ενός ανθρώπου σε ένα κατάστημα, διαφέρει από την κίνηση στον δρόμο. Επίσης, τα χαρακτηριστικά της κίνησης είναι διαφορετικά όταν ένας άνθρωπος απλά περπατάει από όταν περπατάει και μιλάει στο τηλέφωνο ταυτόχρονα.

### **Σχετικά φαινόμενα και συμβάντα**

Οι κινήσεις που εμφανίζονται σε ένα περιβάλλον, επηρεάζονται από τα διάφορα συμβάντα αυτού του περιβάλλοντος. Οι κινήσεις των ανθρώπων επηρεάζονται από το κλίμα και τις καιρικές συνθήκες, από τις αθλητικές και πολιτιστικές εκδηλώσεις, από τις τρέχουσες νομοθετικές

ρυθμίσεις, από τα τέλη διοδίων και τις τιμές των καυσίμων, από τα τροχαία ατυχήματα και ούτω καθεξής. Για την ανίχνευση τέτοιων επιρροών κατά την ανάλυση των δεδομένων, ο αναλυτής πρέπει να χρησιμοποιήσει πρόσθετα στοιχεία και γνώσεις.

## 2.3 Τεχνικές Αναγνώρισης Κίνησης

Η αναγνώριση της ανθρώπινης δραστηριότητας έχει συγκεντρώσει μεγάλο ενδιαφέρον και έχουν πραγματοποιηθεί πολλές έρευνες για τον εντοπισμό πληροφοριών, χρήσιμων για την αναγνώριση κίνησης. Η αναγνώριση κίνησης μπορεί εύκολα να υλοποιηθεί σε κινητά τηλέφωνα χρησιμοποιώντας τους ενσωματωμένους αισθητήρες της συσκευής και γι' αυτό τον λόγο έχουν αναπτυχθεί πολλές εφαρμογές για smartphone. Οι κυριότερες τεχνικές εξόρυξης δεδομένων που χρησιμοποιούν οι υπάρχουσες εφαρμογές για την αναγνώριση της ανθρώπινης δραστηριότητας αναλύονται παρακάτω.

### 2.3.1 Κατηγοριοποίηση (Classification)

Οι περισσότεροι ερευνητές που χρησιμοποιούν την τεχνική της κατηγοριοποίησης για να ανιχνεύσουν την δραστηριότητα του χρήστη, καταγράφουν χαρακτηριστικά (attributes) της κίνησης τα οποία μπορούν να διακριθούν σε τρεις βασικές κατηγορίες. Παρακάτω περιγράφονται τα χαρακτηριστικά που ανήκουν σε κάθε κατηγορία.

- *Μεγέθους.* Αυτή η κατηγορία αναφέρεται σε χαρακτηριστικά που βασίζονται σε τιμές που καταγράφουν οι αισθητήρες της συσκευής. Συνήθως οι τιμές είναι συντεταγμένες, μέση τιμή, τυπική απόκλιση, ελάχιστη και μέγιστη τιμή, μέση απόλυτη διαφορά και άλλες.
- *Συχνότητας.* Είναι τα χαρακτηριστικά που βασίζονται στις τιμές συχνότητας των αισθητήρων. Το πιο συνηθισμένο μέγεθος είναι από τον μετασχηματισμό Φουριερ (ΦΦΤ). Άλλα χαρακτηριστικά αυτής της κατηγορίας είναι η εντροπία της συχνότητας, η μέγιστη συχνότητα, η μέση τιμή, η τυπική απόκλιση και η ενέργεια του μετασχηματισμού Φουριερ.
- *Συσχέτισης.* Πολλοί ερευνητές χρησιμοποιούν ως χαρακτηριστικά, τις συσχετίσεις μεταξύ των υπολοίπων χαρακτηριστικών.

Όσον αφορά στον αλγόριθμο κατηγοριοποίησης που μπορούμε να χρησιμοποιήσουμε για να κατηγοριοποιήσουμε διαφορετικές δραστηριότητες του χρήστη, η επιλογή του εξαρτάται από την

επεξεργαστική ισχύ του συστήματος που θα εκτελεστεί ο αλγόριθμος. Δηλαδή αν ο αλγόριθμος εκτελεστεί σε server είναι προφανές ότι η επεξεργαστική ισχύς του μηχανήματος θα είναι αρκετά πιο μεγάλη από την επεξεργαστική ισχύ ενός smartphone. Οι περισσότεροι ερευνητές χρησιμοποιούν αλγόριθμους κατηγοριοποίησης με επίβλεψη. Αυτοί οι αλγόριθμοι εκπαιδεύονται με κατηγοριοποιημένα δείγματα για τη δημιουργία μοντέλου κατηγοριοποίησης, το οποίο στη συνέχεια θα χρησιμοποιηθεί για την κατηγοριοποίηση των δεδομένων εισόδου. Οι πιο συνηθισμένοι αλγόριθμοι που χρησιμοποιούνται για αναγνώριση κίνησης είναι τα Δέντρα απόφασης (Decision Trees), ο κ - κοντινότερος γείτονας (kNN), ο Naïve Bayes, ο SVM και τα Νευρωνικά δίκτυα. Όμως, η κατηγοριοποίηση με επίβλεψη χρειάζεται μεγάλη επεξεργαστική ισχύ για να δημιουργήσει μοντέλο από δεδομένα εκπαίδευσης και για αυτόν τον λόγο οι περισσότερες υλοποιήσεις έχουν γίνει σε servers. Βέβαια κάποιοι άλλοι ερευνητές, δημιουργούν το μοντέλο κατηγοριοποίησης εκτελώντας τον αλγόριθμο σε ένα μηχάνημα και έπειτα μεταφέρουν το μοντέλο στο τηλέφωνο για την κατηγοριοποίηση των δεδομένων εισόδου. Ακόμα, η αξιολόγηση των αλγορίθμων κατηγοριοποίησης είναι πολύ σημαντική, καθώς δείχνει ποιος αλγόριθμος αποδίδει καλύτερα. Οι δημοφιλέστερες μέθοδοι αξιολόγησης που χρησιμοποιούνται είναι οι n-fold cross validation (συνήθως 10-fold) και τα μέτρα precision, recall, F-measure και accuracy. [3]

### **2.3.2 Συσταδοποίηση τροχιών κίνησης (Spatemporal/Trajectory Clustering)**

Η συσταδοποίηση των τροχιών των κινούμενων αντικειμένων, παρουσιάζει μεγάλο ενδιαφέρον στην ερευνητική κοινότητα και για αυτόν τον λόγο υπάρχει μεγάλος αριθμός ερευνών που εξετάζουν διαφορετικές τεχνικές συσταδοποίησης με σκοπό τη βελτίωση των αποτελεσμάτων τους. Η τροχιά (trajectories) κίνησης ενός χρήστη, αναφέρεται σε μια χρονική ακολουθία τοποθεσιών με τη σειρά που ο χρήστης τις επισκέφτηκε [7]. Η συσταδοποίηση των τροχιών έχει ως στόχο την ομαδοποίηση των τροχιών σε συστάδες όμοιων τροχιών[13]. Οι τροχίες των κινούμενων οντοτήτων πολύ συχνά περιέχουν επαναλαμβανόμενα μοτίβα κίνησης από τα οποία με την τεχνική της συσταδοποίησης είναι δυνατόν να ανιχνευτούν όμοια μοτίβα με σκοπό την αναγνώριση της δραστηριότητας[17].

Οι μέθοδοι που υπάρχουν μέχρι σήμερα στη βιβλιογραφία ακολουθούν δυο βασικές προσεγγίσεις. Η πρώτη προσέγγιση έχει ως στόχο να βρει ένα μέτρο ομοιότητας μεταξύ των τροχιών. Ορίζοντας την απόσταση μεταξύ αντικειμένων καθορίζεται ποιες τροχίες πρέπει να βρίσκονται στην ίδια συστάδα και στη συνέχεια ποιες είναι οι συστάδες που πρέπει να ανιχνευτούν. Ένας βασικός τρόπος για οριστεί η απόσταση είναι να εξεταστούν τροχίες χρηστών που είναι πα-

ρόμοιες, για παράδειγμα τροχιές στις οποίες κάθε χρονική στιγμή οι χρήστες βρίσκονται σχεδόν στην ίδια τοποθεσία. Μια απλή προσέγγιση για τη μοντελοποίηση αυτής της σύγκρισης είναι να θεωρηθούν οι τροχιές διανύσματα των συντεταγμένων και να γίνει σύγκριση των διανυσμάτων χρησιμοποιώντας ένα μέτρο απόστασης, όπως η Ευκλείδεια απόσταση.[8]

Η δεύτερη μέθοδος συσταδοποίησης των τροχιών κίνησης, δεν κάνει χρήση μέτρων ομοιότητας αλλά αντιμετωπίζει κάθε τροχιά ως ένα αντικείμενο και αναζητά ομάδες τροχιών που κινούνται μαζί[8]. Μια αποτελεσματική προσέγγιση χρησιμοποιούν οι Lee και συν. [10], με τον αλγόριθμο TRACCLUS. Αρχικά χωρίζουν τις τροχιές σε μια σειρά υποσυνόλων κάνοντας χρήση του MDL (Minimum Description Length) και έπειτα δημιουργούν συστάδες όμοιων υποσυνόλων λαμβάνοντας υπόψη την πυκνότητά τους. Για κάθε συστάδα, η τροχιά που περιγράφει την συνολική κίνηση των υποσυνόλων που ανήκουν στην ίδια συστάδα, είναι η τροχιά που αντιπροσωπεύει τη συστάδα. Αντίθετα, οι Giannotti και συν. [7] προτείνουν την έννοια των μοτίβων κίνησης και παρουσιάζουν έναν αλγόριθμο για την ανίχνευση τους στις τροχιές. Τα μοτίβα κίνησης αντιπροσωπεύουν σύνολα τοποθεσιών ενδιαφέροντος που σχετίζονται χρονικά, και μπορούν να είναι προκαθορισμένα από τον χρήστη ή να ανακαλυφθούν με κάποιο αλγόριθμο συσταδοποίησης βασισμένο στην πυκνότητα των συστάδων. Ταυτόχρονα, οι Sung, Feldman και Rus[17], παρουσιάζουν έναν αλγόριθμο για την εξαγωγή μοτίβων κίνησης από τις τροχιές και χρησιμοποιούν τα αποτελέσματα για να δημιουργήσουν ένα Μαρκοβιανό μοντέλο το οποίο θα κάνει πρόβλεψη της κίνησης για την κάθε οντότητα.

### **2.3.3 Conditional Random Fields (CRF)**

Μια διαφορετική προσέγγιση για την αναγνώριση δραστηριότητας παρουσιάζουν οι Liao, Lin και Fox κάνοντας χρήση ιεραρχικών μοντέλων CRF σε δεδομένα GPS για την εξαγωγή σημαντικών τοποθεσιών και δραστηριοτήτων του χρήστη.

Αρχικά, χωρίζουν τα δεδομένα GPS ομαδοποιώντας τα με βάση τη σχέση τους στον χώρο και κάθε GPS στίγμα το αντιστοιχούν στον πιο κοντινό δρόμο. Για να δημιουργήσουν μια σωστή συσχέτιση μεταξύ των τοποθεσιών, δημιουργούν ένα μοντέλο CRF με βάση τη χωρική σχέση μεταξύ τους.

Αφού χωρίσουν τις τοποθεσίες, ο αλγόριθμος τους κάνει εκτίμηση για την δραστηριότητα που εκτελείται σε κάθε σύνολο και ξεχωρίζει τις σημαντικές τοποθεσίες κάθε χρήστη. Για να επιτευχθεί αυτό, δημιουργεί ένα νέο μοντέλο CRF το οποίο περιέχει έναν κρυμμένο κόμβο που αντιστοιχεί στη δραστηριότητα σε κάθε σύνολο που δημιουργήθηκε από τις καταγεγραμμένες GPS συντεταγμένες. Κάθε κόμβος δραστηριότητας είναι συνδεδεμένος με μερικά χαρακτηριστι-



κά που προκύπτουν από τις πληροφορίες που προέρχονται από τη διαίρεση των δεδομένων, όπως ημερομηνία και ώρα, μέση ταχύτητα, πληροφορίες για τα σημεία ενδιαφέροντος που βρίσκονται κοντά. Έπειτα, στο μοντέλο CRF που δημιουργήθηκε, εφαρμόζει τον αλγόριθμο Naïve Bayes για να ανιχνεύσει την δραστηριότητα του χρήστη σε κάθε τοποθεσία.

Στη συνέχεια, χρησιμοποιεί τις ακολουθίες στις οποίες ο χρήστης πραγματοποιεί τον ίδιο τύπο κίνησης για να αναζητήσει τις σημαντικές τοποθεσίες του χρήστη. Αυτό γίνεται κατηγοριοποιώντας ξεχωριστές δραστηριότητες στην ίδια ακολουθία, με βάση το αν ανήκουν σε μια συγκεκριμένη τοποθεσία και θεωρώντας ότι όλες οι εγγραφές στις οποίες παρατηρείται μια συγκεκριμένη δραστηριότητα αποτελούν μια σημαντική τοποθεσία.

Η προσέγγιση τους παρουσιάζει ενδιαφέρον, καθώς ο αλγόριθμος που χρησιμοποιούν έχει αρκετά μεγάλο ποσοστό ακρίβειας (86%), ενώ η ταυτόχρονη εκτίμηση δραστηριότητας και τοποθεσίας αυξάνει την ποιότητα του αποτελέσματος. Αυτό συμβαίνει γιατί η κάθε θέση συνδέει τις δραστηριότητες που συμβαίνουν στην χωρική περιοχή της και οι δραστηριότητες αναγνωρίζονται με πιο συνεπή τρόπο.[12]



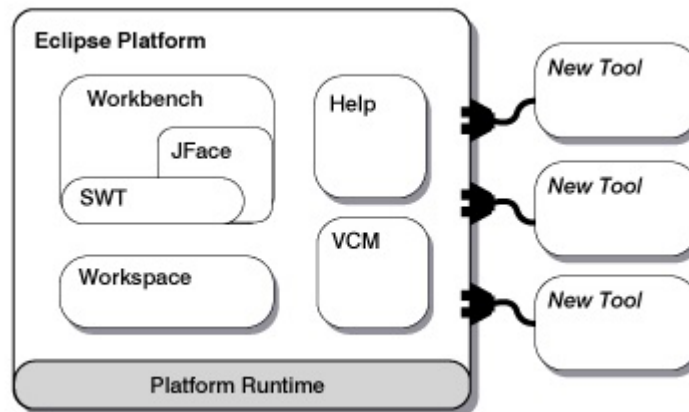
## Κεφάλαιο 3

### Σχεδιασμός

#### 3.1 Το περιβάλλον Eclipse

Το Eclipse IDE είναι ένα περιβάλλον ανάπτυξης λογισμικού, το οποίο αρχικά αναπτύχθηκε από την IBM ως εργαλείο ανάπτυξης Java με σκοπό να αντικαταστήσει το ήδη υπάρχον περιβάλλον Visual Age, αλλά κυκλοφόρησε ως λογισμικό ανοιχτού κώδικα τον Νοέμβριο του 2001[4]. Το 2004, ο μη κερδοσκοπικός οργανισμός Eclipse Foundation και το επιστημονικό του προσωπικό ανέλαβε ολοκληρωτικά τον έλεγχο της πλατφόρμας Eclipse. Το Eclipse σήμερα αποτελεί το πιο διαδεδομένο περιβάλλον ανάπτυξης για Java, ενώ ταυτόχρονα εξαιτίας της επεκτάσιμης αρχιτεκτονικής του χρησιμοποιείται ως εργαλείο ανάπτυξης για πολλές άλλες γλώσσες προγραμματισμού[5].

Η πλατφόρμα Eclipse σχεδιάστηκε με στόχο την ανάπτυξη ολοκληρωμένων περιβαλλόντων ανάπτυξης. Αποτελεί ένα εύκολα επεκτάσιμο περιβάλλον, παρέχοντας μηχανισμούς που επιτρέπουν την εγκατάσταση επιπρόσθετων εργαλείων τα οποία μπορούν να κάνουν χρήση άλλων διεπαφών ανάπτυξης εφαρμογών (APIs). Η αρχιτεκτονική της πλατφόρμας Eclipse είναι δομημένη με τέτοιο τρόπο ώστε να επιτρέπει την χρήση πρόσθετων εργαλείων, όπως φαίνεται στο Σχήμα 3.1[5].



Σχήμα 3.1: Η αρχιτεκτονική της πλατφόρμας Eclipse

Τα πρόσθετα εργαλεία αποτελούν την μικρότερη μονάδα της πλατφόρμας Eclipse, τα οποία είναι δομημένα κομμάτια κώδικα που προσφέρουν επιπλέον λειτουργικότητα στην πλατφόρμα. Για την ανάπτυξη μιας εφαρμογής για Android, χρησιμοποιείται ένα σύνολο πρόσθετων που είναι γνωστό ως Android Development Toolkit (ADT) το οποίο επεκτείνει τα υπάρχοντα εργαλεία της Java (JDK) για να προσφέρει συγκεκριμένες λειτουργίες που απαιτούνται για τον προγραμματισμό σε Android.

## 3.2 Το ADT plug-in

Το ADT (Android Developer Tools) είναι ένα πρόσθετο για το Eclipse το οποίο περιλαμβάνει ένα σύνολο εργαλείων που ενσωματώνονται στο περιβάλλον του Eclipse. Προσφέρει πρόσβαση σε πολλές λειτουργίες που διευκολύνουν την ανάπτυξη εφαρμογών Android. Το ADT παρέχει γραφικό περιβάλλον για την πρόσβαση στα εργαλεία του SDK καθώς και ένα εργαλείο σχεδιασμού γραφικού περιβάλλοντος για την εύκολη και γρήγορη δημιουργία της διεπαφής χρήστη της εφαρμογής που αναπτύσσεται.

Η εγκατάσταση του ADT διευκολύνει τη δημιουργία και την δοκιμή εφαρμογών Android, επιτρέποντας την δημιουργία, την ανάπτυξη και τον εντοπισμό σφαλμάτων της εφαρμογής ενώ επεκτείνει το documentation της Java για τα APIs του Android. Ακόμα, ενσωματώνει τα εργαλεία του Android SDK, προσθέτοντας τα αντίστοιχα μενού στο περιβάλλον του Eclipse. Επιπλέον, το ADT συμπεριλαμβάνει επεξεργαστή XML αρχείων, το οποίο επιτρέπει την επεξεργασία των XML αρχείων του Android και τη δημιουργία γραφικών διεπαφών χρήστη.

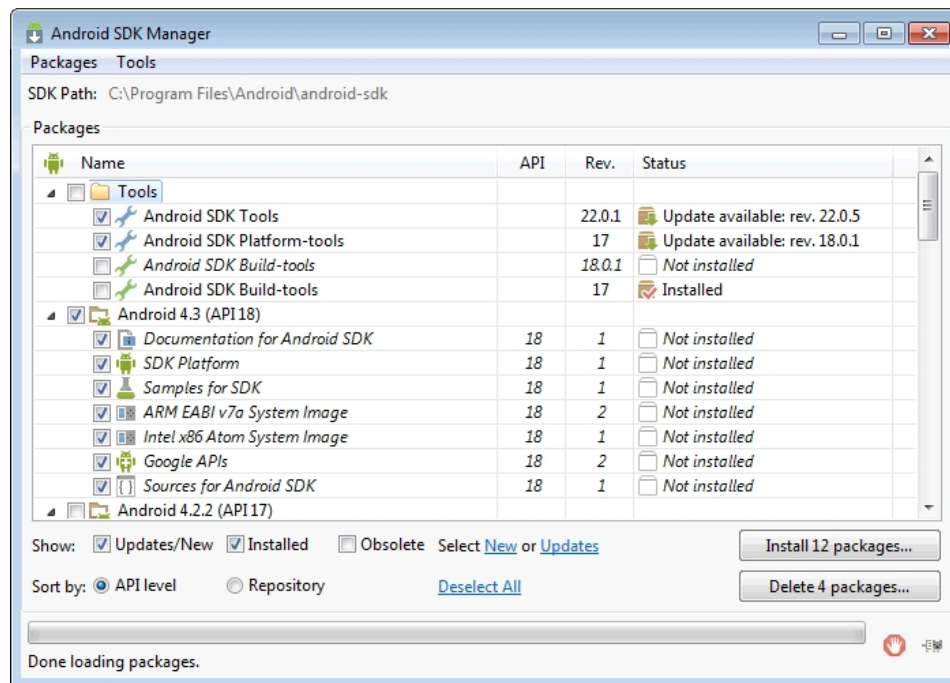
### 3.3 Το εργαλείο Android SDK

Το Android SDK (Software Development Kit) περιλαμβάνει ένα σύνολο εργαλείων, τα οποία είναι απαραίτητα για την ανάπτυξη εφαρμογών για συσκευές Android και ενσωματώνεται στο περιβάλλον Eclipse μέσω του ADT. Τα εργαλεία διακρίνονται σε δυο κατηγορίες: εργαλεία SDK και εργαλεία πλατφόρμας Android. Τα εργαλεία SDK είναι ανεξάρτητα από την έκδοση Android για την οποία προορίζεται η εφαρμογή, ενώ τα εργαλεία πλατφόρμας προσαρμόζονται ώστε να υποστηρίζουν τα χαρακτηριστικά του κάθε API Level που χρησιμοποιείται.

#### Τα εργαλεία SDK

Η χρήση των εργαλείων SDK είναι απαραίτητη για την ανάπτυξη εφαρμογών Android στο περιβάλλον Eclipse, τα πιο σημαντικά εργαλεία περιγράφονται παρακάτω.

##### 3.3.1 SDK Manager



Σχήμα 3.2: Στιγμιότυπο του SDK Manager.

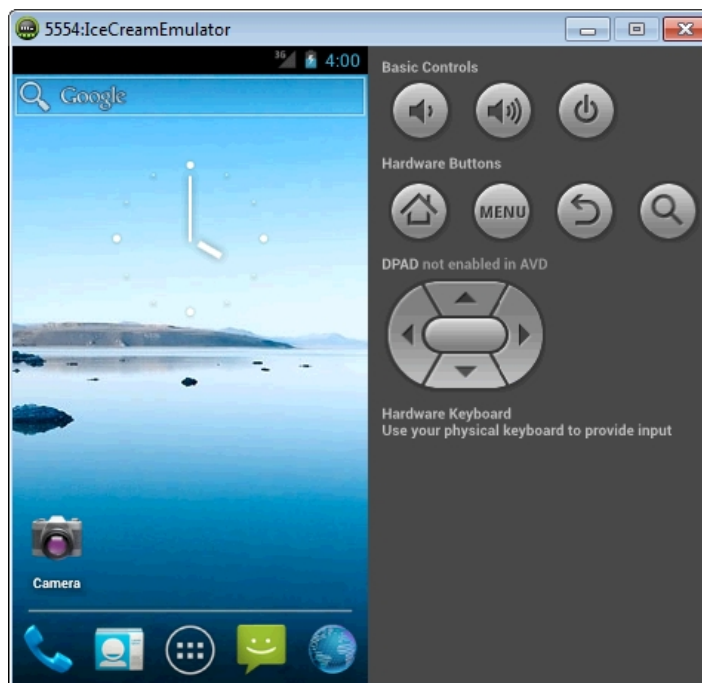
Ο SDK Manager, διαχειρίζεται το Android SDK χωρίζοντας τα εργαλεία και τις πλατφόρμες σε πακέτα και επιτρέπει στον χρήστη να κατεβάσει τα πακέτα που του είναι απαραίτητα. Ο SDK

Συλλογή και κατηγοριοποίηση δεδομένων κίνησης, μέσω κινητών τηλεφώνων, με στόχο την ανίχνευση δραστηριότητας

---

Manager εμφανίζει τα πακέτα SDK που είναι διαθέσιμα, όσα έχουν ήδη εγκατασταθεί αλλά και αυτά για τα οποία υπάρχουν διαθέσιμες ενημερώσεις όπως φαίνεται στο Σχήμα 3.2.

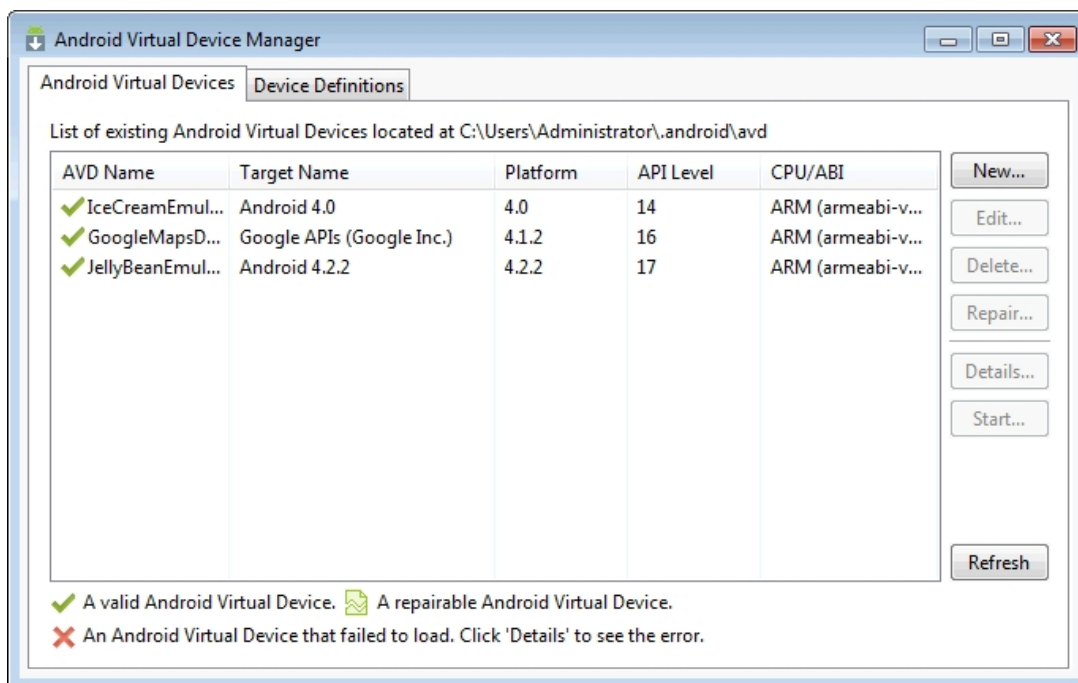
### 3.3.2 Προσομοιωτής κινητής συσκευής (Android Emulator)



Σχήμα 3.3: Στιγμιότυπο του Android Emulator

Το Android SDK περιλαμβάνει έναν προσομοιωτή κινητής συσκευής, ο οποίος τρέχει στον υπολογιστή του χρήστη και του επιτρέπει να αναπτύσσει και να δοκιμάζει εφαρμογές Android χωρίς τη χρήση φυσικής συσκευής.

### 3.3.3 AVD Manager

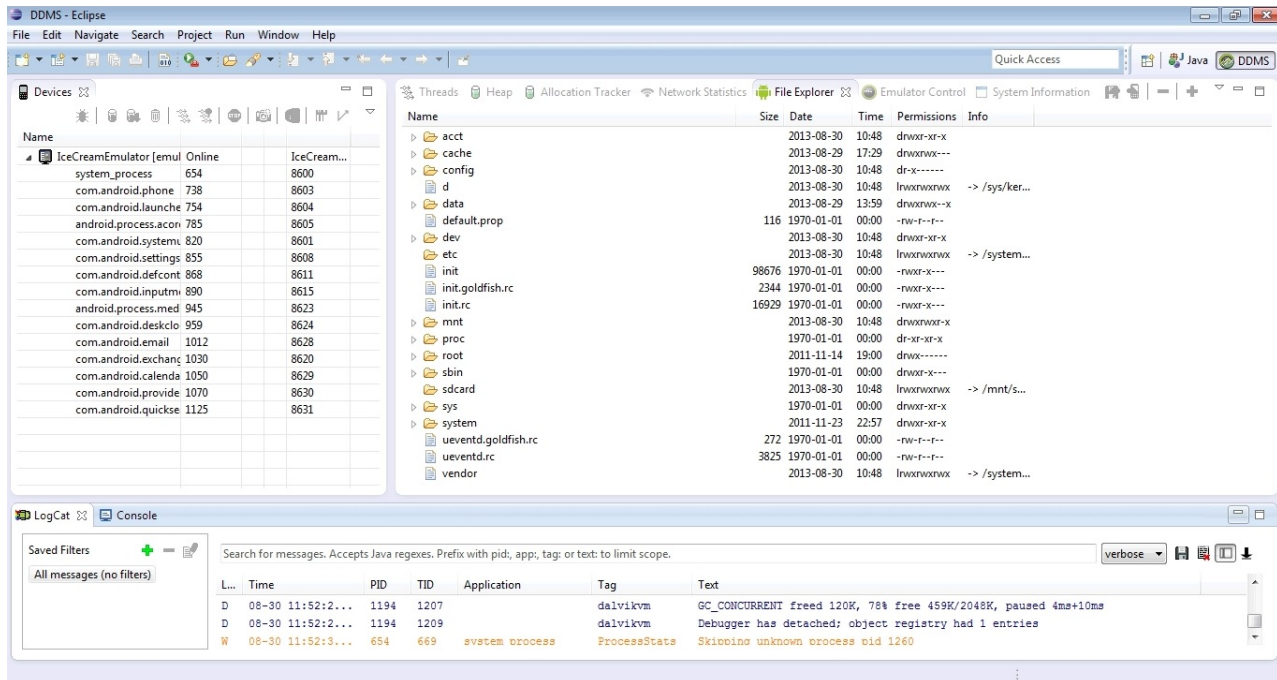


Σχήμα 3.4: Στιγμιότυπο του AVD Manager.

Ο AVD (Android Virtual Device) Manager αποτελεί ένα γραφικό περιβάλλον για τη διαχείριση των ρυθμίσεων των εικονικών συσκευών Android. Μια συσκευή AVD ορίζει τις ρυθμίσεις διαμόρφωσης του προσομοιωτή συσκευής Android επιτρέποντας την αναπαράσταση διαφορετικών συσκευών που τρέχουν λειτουργικό σύστημα Android. Ο AVD Manager, επιτρέπει τη δημιουργία, τη διαγραφή και την επισκευή των AVD αλλά και την εμφάνιση των ρυθμίσεων για κάθε συσκευή.

Συλλογή και κατηγοριοποίηση δεδομένων κίνησης, μέσω κινητών τηλεφώνων, με στόχο την ανίχνευση δραστηριότητας

### 3.3.4 DDMS (Dalvik Debug Monitor Server)



Σχήμα 3.5: Στιγμιότυπο του DDMS.

Το DDMS είναι ένα εργαλείο για debugging το οποίο περιλαμβάνεται στο Android SDK και τρέχει είτε στον emulator είτε σε φυσική κινητή συσκευή. Παρέχει υπηρεσίες port-forwarding και καταγραφής συμβάντων (logcat), καταγράφει στιγμιότυπα της οθόνης της συσκευής, εμφανίζει πληροφορίες για τα νήματα και τον σωρό της συσκευής ενώ προσομοιώνει εισερχόμενες κλήσεις, SMS και λήψη σήματος GPS.

## Τα εργαλεία πλατφόρμας Android

### 3.3.5 ADB (Android Debug Bridge)

Το ADB είναι ένα εργαλείο γραμμής εντολών το οποίο επιτρέπει το debugging του Android κώδικα μέσω του Eclipse. Το DDMS και το ADT χρησιμοποιούν το ADB για να διευκολύνουν την επικοινωνία του περιβάλλοντος ανάπτυξης και της συσκευής. Ακόμα, το ADB χρησιμοποιείται για την πρόσβαση στο σύστημα αρχείων της συσκευής, την χειροκίνητη εγκατάσταση και την απεγκατάσταση εφαρμογών Android στη συσκευή και την εκτέλεση εντολών φλοιού.



### **3.4 Ανάπτυξη Android Εφαρμογής στο περιβάλλον Eclipse**

### **3.5 Το περιβάλλον Weka**

#### **3.5.1 Χρήση του Weka με κώδικα Java**

### **3.6 Οι χάρτες Google σε περιβάλλον Android**



## **Κεφάλαιο 4**

### **Υλοποίηση**



## **Κεφάλαιο 5**

### **Αποτελέσματα**



## **Κεφάλαιο 6**

### **Συμπεράσματα**





# Βιβλιογραφία

- [1] Μαρία Χαλκίδη Μιχάλης Βαζιργιάννης. *Εξόρυξη Γνώσης από Βάσεις Δεδομένων και του Παγκόσμιο Ιστό*. ΤΥΠΩΘΗΤΩ - Γιώργος Δαρδάνος, 2005.
- [2] Hagen Höpfner. Smartphone Hardware Sensors. [Online; τελευταία προσπέλαση 11-Αυγούστου-2013].
- [3] Mohd Fikri Azli bin Abdullah, Ali Fahmi Perwira Negara, Md Shohel Sayeed, Deok-Jai Choi, and Kalaiarasi Sonai Muthu. classification algorithms in human activity recognition using smartphones. *Proceedings of World Academy of Science, Engineering and Technology*, 68, 2012.
- [4] Ed Burnette. *Eclipse IDE Pocket Guide*. O'Reilly Media, Inc., 2005.
- [5] Onur Cinar. *Android Apps with Eclipse*. Apress, Berkely, CA, USA, 1st edition, 2012.
- [6] William Enck, Peter Gilbert, Byung-Gon Chun, Landon P. Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N. Sheth. taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In *Proceedings of the 9th USENIX conference on Operating systems design and implementation*, OSDI'10, pages 1-6, Berkeley, CA, USA, 2010. USENIX Association.
- [7] Fosca Giannotti, Mirco Nanni, Dino Pedreschi, Fabio Pinelli, Chiara Renso, Salvatore Rinzivillo, and Roberto Trasarti. mobility data mining: discovering movement patterns from trajectory data. In *Proceedings of the Second International Workshop on Computational Transportation Science*, IWCTS '10, pages 7-10. ACM, 2010.
- [8] Gianni Giannotti, Fosca Giannotti, and Dino Pedreschi. *Mobility, data mining and privacy: Geographic knowledge discovery*. Springer, 2008.

- [9] Paolo Giudici and Silvia Figini. *Applied Data Mining for Business and Industry*. Wiley Publishing, 2nd edition, 2009.
- [10] Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. trajectory clustering: a partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, SIGMOD '07, pages 593–604. ACM, 2007.
- [11] Wei-Meng Lee. *Beginning Android Application Development*. Wrox Press Ltd., Birmingham, UK, UK, 1st edition, 2011.
- [12] Lin Liao, Dieter Fox, and Henry Kautz. extracting places and activities from gps traces using hierarchical conditional random fields. *Int. J. Rob. Res.*, 26:119–134, 2007.
- [13] Gerasimos Marketos and Y Theodoridis. mobility data warehousing and mining. *Proceedings of VLDB PhD Workshop*, 2009.
- [14] Matthias Steinbauer, Ismail Khalil, and Gabriele Kotsis. reality mining at the convergence of cloud computing and mobile computing. *ERCIM News*, 2013.
- [15] Mei-po Kwan, Jiyeong Lee, In Michael, F. Goodchild, and Donald G. Janelle. geovisualization of human activity patterns using 3d gis: A time-geographic approach. In *Spatially Integrated Social Science*, pages 48–66. University Press, 2004.
- [16] Sweta Pittie, Hillol Kargupta, and Byung-Hoon Park. dependency detection in mobimine: a systems perspective. *Inf. Sci. Inf. Comput. Sci.*, 155(3-4):227–243, 2003.
- [17] Cynthia Sung, Dan Feldman, and Daniela Rus. trajectory clustering for motion prediction. In *IROS*, pages 1547–1552. IEEE, 2012.
- [18] Domenico Talia and Paolo Trunfio. *Mobile Data Mining on Small Devices through Web Services*, pages 264–276. John Wiley & Sons, Inc., 2010.
- [19] Developers Android. Android Platform Versions Dashboard. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].
- [20] Wikipedia, The Free Encyclopedia. Accelerometer. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].
- [21] Wikipedia, The Free Encyclopedia. Android(operating system). [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].

- [22] Wikipedia, The Free Encyclopedia. Global Positioning System. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].
- [23] Wikipedia, The Free Encyclopedia. GPS navigation device. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].
- [24] Wikipedia, The Free Encyclopedia. Magnetometer. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].
- [25] Wikipedia, The Free Encyclopedia. Smartphone. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].
- [26] The University of Waikato. Use WEKA in your Java code. [Online· τελευταία προσπέλαση 11-Αυγούστου-2013].