View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Oulu Repository - Jultika

UNIVERSITY OF OULU

FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Junjie Zhou

A Mobile Sensing Solution for Indoor and Outdoor State Detection

Master's Thesis
Degree Programme in Computer Science and Engineering
Nov 2016

Zhou J. (2016) A Mobile Sensing Approach for Indoor and Outdoor State Detection. University of Oulu, Degree Programme in Computer Science and Engineering. Master's Thesis, 32 p.

ABSTRACT

One important research challenge in ubiquitous computing is determining a device's indoor/outdoor environmental state. Particularly with modern smartphones, environmental information is important for enabling of new types of services and optimizing already existing functionalities.

This thesis presents a tool for Android-powered smartphones called ContextIO for detecting the device's indoor/outdoor state by combining different onboard sensors of the device itself. To develop ContextIO, we developed a plugin to AWARE mobile sensing framework. Together the plugin and its separate controller component collect rich environmental sensor data. The data analysis and ContextIO's design considers collected data particularly about magnetometer, ambient light and GSM cellular signal strength. We manually derive thresholds in the data that can be used in combination to infer whether a device is indoor or outdoor. ContextIO uses the same thresholds to infer the state in real time.

This thesis contributes an Android tool for inferring the device's indoor/outdoor status, an open dataset that other researchers can use in their work and an analysis of the collected sensor data for environmental indoor/outdoor state detection.

Keywords: ubiquitous computing, smartphones, mobile sensing, positioning

Zhou J. (2016) Mobiilihavainnointiin perustuva ratkaisu älylaitteen paikantamiseksi sisä- ja ulkotilojen suhteen. Oulun yliopisto, tietotekniikan tutkinto-ohjelma. Diplomityö, 32 s.

TIIVISTELMÄ

Yksi jokapaikan tietotekniikan tutkimuskysymyksistä keskittyy selvittämään onko laitteen sijainti sisä- vai ulkotilassa. Etenkin uudet älypuhelimet pystyvät hyödyntämään tätä tietoa uudenlaisten palveluiden ja sovellusten kehittämisessä sekä vanhojen toiminnallisuuksien optimoinnissa.

Tämä diplomityö esittelee Android-käyttöjärjestelmällä toimiville puhelimille suunnatun työkalun nimeltään ContextIO. Työkalu yhdistelee älypuhelimen sensorien tuottamaa tietoa ja havaitsee laitteen siirtymisen eri sijaintiin sisä- ja ulkotilojen suhteen. ContextIO:n suunnittelu ja kehitystyö perustuvat data-analyysiin, jonka data kerättiin AWARE-sensorialustan liitännäisellä sekä erillisellä nimeämistyökalulla. Data-analyysi keskittyy magnetometrin, valosensorin sekä GSM-kentän voimakkuuden hyödyntämiseen paikantamisessa. Kerätystä datasta määriteltiin raja-arvot, joita yhdistelemällä voidaan varsin luotettavasti todeta laitteen sijainti sisä- ja ulkotilojen suhteen. Nämä raja-arvot luovat perustan ContextIO:n reaaliaikaiselle laitteen sijainnin määrittämiselle.

Tämän diplomityön pääasialliset tulokset ovat työkalu Android-pohjaisten älypuhelinten sijainnin määrittämiseen sisä- ja ulkotilojen suhteen, avoin datasetti, jota muut tutkijat voivat käyttää sekä sijainnin määrittämiseen keskittyvä data-analyysi.

Avainsanat: jokapaikan tietotekniikka, älypuhelimet, mobiilihavainnointi, sijainnin määrittäminen

TABLE OF CONTENTS

FOREWORD

This thesis was completed for Center for Ubiquitous Computing at the University of Oulu, Finland. First, I would like to give my regards to my Supervisor, Dr. Simo Hosio for spending a great amount of time in assisting the thesis work in both the programming phase and the documenting phase. In addition, I want to extend my regards to my second supervisor, Dr. Jorge Goncalves, for his advice in improving the thesis. I would also like to thank the whole Center for Ubiquitous Computing for providing the devices to develop and test the software. I thank Mr. Chu Luo for providing inspiration on how to improve developed software in the future.

Second, I would also like to thank my friends, Pingjiang, Kai, Zeyun, Jilin, and Haejong for their accompany and help during the 2 years' study period. Besides, I am grateful to the host country, Finland, for proving good and free chance for education.

Finally, yet most importantly, I want to deeply express my gratitude to my family members, parents and grandparents for endless love and encouragement.

Oulu, 20.11.2016

Junjie Zhou

1. INTRODUCTION

The availability of accurate contextual information plays an increasingly important role in current mobile smartphone platforms, both for optimizing old and enabling new types of applications. Context-aware mobile applications that utilise mobile sensing techniques can, for example, coordinate and alter different settings based on context information such as ambient noise, phone profile or weather [1, 2]. Or, if a smartphone is aware of its indoor/outdoor state it can adjust volume and vibration settings of notifications accordingly.

Several projects and prototypes are seeking to accurately detect a device's indoor/outdoor state and solutions such as location-based labeling or using simple sensor data have been proposed. Still, there is a clear need for alternative efficient and lightweight solutions. One of the promising ways for determining a phone's environmental state is fingerprinting ambient information [3] that includes both physical info such as sound, motion, colour, and logical information such as mapping WiFi networks and other types of mobile signal strengths. After constructing a rich fingerprint database about the surroundings, detecting the state becomes a matter of database queries and comparing the real time sensor readings to the stored fingerprints and choosing the best fit. However, this technique naturally requires a number of sensors and a vast database. Further, ambient noise and availability of WiFi networks tend to change fast, making this technique suboptimal.

The availability of onboard smartphone sensors is constantly increasing. This makes real time mobile sensing a powerful enabler for new kinds of applications. New types of games and services, such as *Pothole Patrol* that uses sensors to monitor the surface conditions of roads [4] or *BikeNet Mobile Sensing System* that map cycling experiences [5], are gaining popularity.

In this thesis, we leverage real time mobile sensing to detect environmental indoor/outdoor state of the device. We focus on and discuss several topical issues such as energy efficiency, detection accuracy and the availability of sensors in the majority of Android phones on the market at the time of writing this thesis.

1.1. Objective and Research Questions

The objective of this thesis is to build an Android application to detect indoor/outdoor state of the device. To support the objective, this thesis focuses on two research questions:

RQ1: What sensors are available and can be used for facilitating indoor/outdoor state detection on a modern Android smartphone?

RQ2: More specifically, which factors and sensors influence most the accuracy of indoor/outdoor state detection and should therefore be included in the final tool?

1.2. Structure of this Thesis

Chapter 2 introduces mobile sensing and indoor/outdoor state detection solutions. Chapter 3 presents data collection and analysis. The implemented solution is then introduced in Chapter 4. Chapter 5 discusses the work in the context of related work and presents a summary of findings and the author's experiences.

2. MOBILE SENSING

This chapter briefly introduces mobile sensing in general, and then proceeds to review already existing solutions for inferring indoor/outdoor state with emphasis on mobile computing and smart phones.

2.1. Introduction

Mobile phones have rapidly developed into powerful sensing platforms, and new capabilities are increasingly being used to sense everyday events and physical surroundings to enable new types of applications and features [6]. For example, devices' motion (accelerometer, gyroscope), location or ambient noise can already be measured using off-the-shelf smartphones [7]. However, mobile sensing introduces a number of challenges as well, such as accuracy of commonly used sensors (e.g. GPS location, barometer), energy efficiency issues and efficient mining and interpretation of collected data [6].

However, applications using mobile sensing are already mature enough for commercial distribution, and mobile application markets such as Google Play and Apple App Store contain a plethora of mobile sensing applications for a variety of different purposes [8].

2.2. Context

A broad generalization can be made that all mobile sensing aims at inferring different contexts. Context, according to Cambridge English dictionary, means "the situation within which something exists or happens" [9], and this is also the definition we use in this thesis. However, context has a variety of different descriptions, and different researchers have different understanding towards context. For instance, when describing the concept of context-aware, context has been defined as location of use, collection of nearby objects, such as people, host, and accessible devices and changes to such objects over time [2, 10]. Chen and Kotz categorized context into 4 different groups, namely computing context, user context, physical context and time context [11]. Dey and Abowd denoted context as any information that can be used to characterize the situation of an entity and also categorize the context as who, what, when, where, why [12].

2.3. Smartphone sensor capabilities

While the early cellphones contained hardware merely for supporting communication, manufacturers rapidly realised the potential of adding value to the device with cameras, compasses, barometers, proximity sensors, light sensors, GPS functionalities, etc. In early 2000s, the so called "smartphones" started to globally proliferate. Today's smartphones are powerful networked computing devices that often are used for much more than just communication purposes. The modern smartphone can truly sense, see and hear ambient environments [13].

In this thesis, we mainly focus on Android-powered devices, as in 2016 almost 9 out of 10 sold smartphones feature Android OS [14]. The onboard sensors in Android smartphones are classified into 3 broad categories, namely motion sensors, environmental sensors and position sensors [15]. Motion sensors include such as accelerometers, gravity sensors, gyroscopes, and rotational vector sensors. Environmental sensors include barometers, light sensor (luminance), cameras, thermometers, humidity, etc. [15]. Finally, position sensors aim to help physically locating the device using for instance GPS, magnetometer, and compass.

2.4. Aggregating sensor data

The data produced by onboard sensors can also be combined. Aggregating different types of data can produce more accurate results especially for inferring information about different contexts than using a single data source. Similarly, aggregated sensor data is commonly utilised in researching human behavior and social activities [16]. However, this has been shown leads to issues with energy consumption and privacy [16].

Internet of Things (IoT) broadly refers to enabling networking between various types of computing devices [17]. As the number of networked devices increase, so does the need for coordinating and managing data. Several modern platforms, such as *PubNub* [18], aim to synchronize data that flows between devices and the cloud using a publish/subscribe model. Another example specifically for mobile devices is *Mobile Sensor Data Collector* that synchronizes data via Bluetooth and WiFi [19]. The requirement for real time data transfer in such platforms has raised a need for considering the transfer protocols carefully. For example, *Mobile Sensor Data Collector* [19] chooses UDP over TCP protocol in order to avoid confirmation signals. To reduce the energy consumption, different platforms take different approaches. For example, *Piggyback CrowdSensing* [20] intelligently leverages the period when user is operating the smartphone to collect sensor data. This way, since many smartphone resources, such as CPU or other sensors, are activated already, energy consumption can be reduced significantly. *LittleRock* [21], when possible, offloads the sampling and processing of sensor data to low-power processors instead of the main processor, and thus enables the phone to perform continuous sensing at a low power overhead [21].

To tackle all these issues and enable easy aggregation of sensor data, several mobile sensing frameworks exist. Next, we will introduce some of the frameworks that are most related to our approach.

2.5. Mobile sensing frameworks

There exist several mobile sensing frameworks that make it easy to aggregate sensor data. Here, we first discuss AWARE, as the most related framework to the thesis, and continue with two more examples of mobile sensing frameworks.

Aware framework [22] is a mobile instrumentation framework that enables easy data collection and aggregation. It is aimed at *"facilitating our understanding of human behavior."* Specifically, AWARE can be utilized to build context-aware applications, collect data, and study human behavior. AWARE offers a configurable mobile client that handles sensor data collection using a variety of different onboard sensors and software sensors. To handle more complicated data collection needs, AWARE can also be embedded in another application and it also supports plugins that can implement custom functionality and use AWARE's sensing capabilities. The sensor data is timestamped and first collected in a local SQLite database on the phones. Then, a sync interval defines when the data is synced to a server, also offered by AWARE for researchers to store and retrieve user data. Ultimately, AWARE is a flexible framework that is useable for a wide range of purposes. Thus, several academic studies that focus on studying human behavior have used AWARE as the data collection platform (see e.g., [23] or [24]).

MyExperience is a mobile middleware dedicated to understanding the user's perception, motivation and satisfaction of mobile device by means of capturing both subjective and objective data (i.e., human-based and sensor-based data) **[25]**. The subjective data, namely human-based data, is collected through a context-triggered Experience Sampling Method (ESM) survey. MyExperience targets two specific group of audiences, namely researchers and participants. Researchers customize the tool according to their specific study requirements, and participants run the tool in various capacities based on the customizations made by researchers **[25]**. As with AWARE, data collected by MyExperience can be synchronized to a server in near real time.

Funf [26] is a middleware focusing on social interaction and behavior. Funf collects data from 25 phone-based sources of signal which can primarily be classified as hardware sensors (e.g., GPS, accelerometer, Bluetooth), software sensors (e.g., calls, SMS), communication activities, installed and running applications, multimedia information, manually input data by participants, etc. Data is collected through 3 approaches, namely *Mobile Phone Sensing Platform*, *Surveys*, and *Facebook Data Collection Application*. Specifically, Mobile Phone Sensing Platform is the core of data collection. This platform runs in a phone periodically, and it senses and records information from all the 25 types of data signals. Survey has 2 types, namely web-based and on-phone surveys. Survey, supposed to be completed at regular intervals, has monthly surveys and daily surveys. Facebook Data Collection Application is an option for users to install on their own. It logs users' online activities. The phone data is stored in local storage with a SQLite file format. Every 3 hours data file is changed to reduce data loss due to file corruption and periodic data uploading to back-end is also supported.

2.6. Approaches to detecting indoor/outdoor state

Typically, in literature about indoor/outdoor state detection, the environment is classified in three categories: indoor, semi-outdoor and outdoor. Specifically, indoor usually refers to being inside a building, and semi-outdoor refers to being close to a building or in a semi-open building. Outdoor refers to being outside any building. Semi-outdoor is conceptually the most ambiguous, but the need for it is justified, as it clearly differs from the others for example in the typical availability of wireless networks [27].

One popular means to explore environmental states is image processing and pattern recognition [28, 29, 30]. Such technique primarily uses camera and light sensors, which are already available in off-the shelf smartphones.

One of the most straightforward methods for indoor/outdoor detection is leveraging the availability of the GPS signal. Outdoor, GPS is typically available whereas indoor the signal strength is either significantly weaker or not detectable at all [31]. However, in some cases, for example when the device is close to an open window or door, GPS can receive adequate signal, which reduces its reliability for state detection [31]. In addition, GPS plays an role of high energy consumption sensor among the whole sensor collection [32], thus GPS is not a suitable choice in terms of continuous sensing demand. Eventually GPS is not chosen to be utilized.

Radio Frequency Identification (RFID) including both active [33, 34] and passive types [35] can be used to facilitate indoor/outdoor positioning. In this case developers program the Radio Frequency (RF) tags that are used to infer the position of the user. In case of RFID, the tag reader can be either embedded in the environment or the device being positioned can carry the reader and the tags are placed in the surroundings. Similar to RFID, other positioning systems also rely on custom wireless signal infrastructure, and examples of such systems can be found in [36, 37].

A common method for fingerprinting different locations, and therefore also enabling indoor/outdoor inference, is utilising received signal strength (RSS). Examples of systems and methods can be found in [38, 39, 40, 41, 42, 43, 44]. Different networks can be utilised in fingerprinting such as WiFi RSS [38, 39, 40, 41, 42, 43], cellular network [44] and custom third party transmitters [36]. Of these, WiFi now is practically ubiquitous inside buildings in many parts of the world [39] and is therefore perhaps the most robust choice.

Finally, magnetic field variation [45] and light intensity [27, 46] have been also successfully leveraged to infer indoor/outdoor environmental state.

2.7. Mobile indoor/outdoor state detection solutions

Several mobile solutions that use one or a combination of the previously introduced technological enablers to infer the device's indoor/outdoor state exist. For instance, *IODetector* is an Android application that uses leverages onboard ambient light sensor, magnetometer and cellular signal strength and time of day to detect between indoor, outdoor and semi-outdoor contexts [27]. *IODetector* offers a user interface for users to input ground truth about the context on the go, which is then compared to the inferred state to demonstrate accuracy. *IODetector* uses only energy-efficient sensors, and thus is fairly "lightweight".

IODetector uses static thresholds in the sensed data for state detection. In certain conditions (*e.g.* geographic areas) this leads to problems unless the thresholds are manually changed. An alternative solution to static thresholds is using semi-supervised machine learning using Co-Training method that produces a more robust result in changing environments (see examples in [31, 46). Semi-Supervised machine learning refers to training the final classifier (here, the "decision-maker" about the context) with data that is not fully labeled, *i.e.,* there is not always ground truth available. Further, co-Training [31] method focuses on generating new data set in new environment and generates new model based on the newly merged data set. Specifically, co-training refers to using 2 classifiers to train data collaboratively and learn from each other. However, this approach is battery and CPU intensive.

Using the current mobile devices, it is still unrealistic to continuously train classifiers and simultaneously collect and store data from several sensors. For this, battery and CPU technologies must develop significantly more. Another approach to machine learning is using GPS positioning combined with sensor-based indoor/outdoor detection [47]. Specifically, the sensor-based solution infers indoor/outdoor state, while GPS is opened outdoor to find the transition location

[48]. If indoor, the device attempts to use WiFi signal strength to find the transition location. However, this solution is also fairly battery and CPU intensive.

3. DATA COLLECTION AND ANALYSIS

Here we present the means and procedure for collecting rich environmental context data. The data is then analysed and findings used to inform the design of our own mobile indoor/outdoor detection solution. This thesis uses three contextual states: indoor, semi-outdoor and outdoor. The states and their meaning in our context are introduced in more detail later.

3.1. Mobile Toolkit

The data collection relies on two components, namely a plugin for the AWARE mobile instrumentation framework discussed earlier [22, 49] and a remote-control application on a different device. When using AWARE, researchers who want to collect data must first create a *study* by using an online portal provided by AWARE. There, researchers simply indicate which sensor data should be collected and how frequently should a sample of each type be stored. Then, AWARE devices that join the study will automatically be configured accordingly. However, in this case raw sensor data is not enough, as we must also know where do the samples originate. More specifically, we wish to know if the samples are from indoor, semi-outdoor or outdoor environments.

To enable ground truth state collection, we implemented an AWARE plugin that was used to accept ground truth data about environmental state (indoor/semi-outdoor/outdoor) via Bluetooth. After startup, it simply waits in a server mode for a Bluetooth connection to be made. To accept an inbound Bluetooth connection to the plugin the user has to perform typical Bluetooth operations such as enabling discovery and pairing the device with a potential connecting device.

Then, the remote labeler is used to connect to the plugin, and to relay state transition information. The labeler features buttons for starting and stopping a data collection session and for each possible transition between indoor, semi-outdoor and outdoor states. Figure 1 clarifies the Remote Labelers user interface.

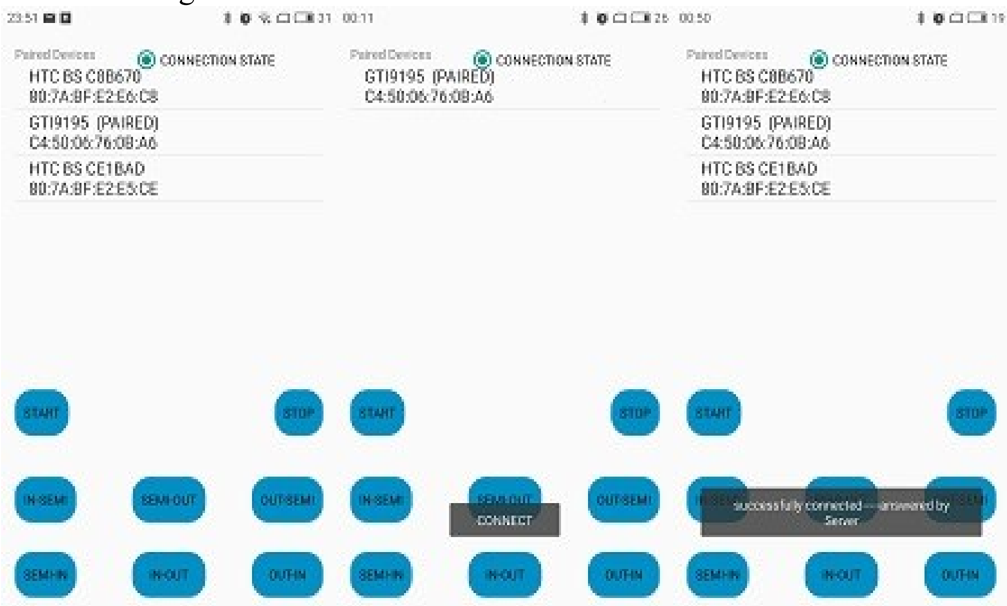


Figure 1. From left: the Remote Labeler displays all discoverable devices, the connection is being attempted and the connection is accepted by the AWARE plugin.

In practice, any user wishing to collect data first creates a connection to the AWARE plugin, pushes “start” to indicate start of data collection. Then, whenever the user goes into a different environmental state, for example from indoor to outdoor, she must use the Remote Labeler’s respective button to send a signal to the AWARE plugin about state transition. The plugin then stores the transition timestamp and labels the previously collected sensor data. Naturally, at the same time it knows from which state is the next data batch originating from. Figure 2 further illustrates the process between the two components.

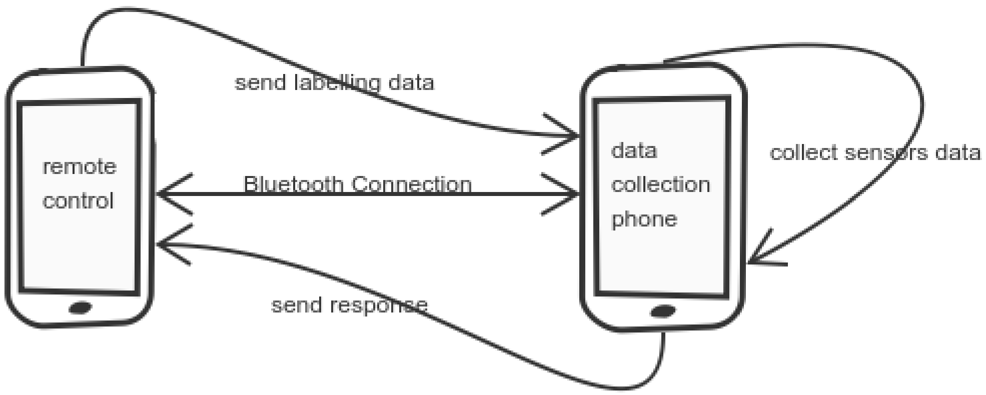


Figure 2. Conceptual diagram of the mobile solution to collect data for analysis.

The AWARE plugin offers a minimal user interface towards the user collecting data, as all needed interaction happens with the remote labeler. Figure 3 depicts the plugin interface.

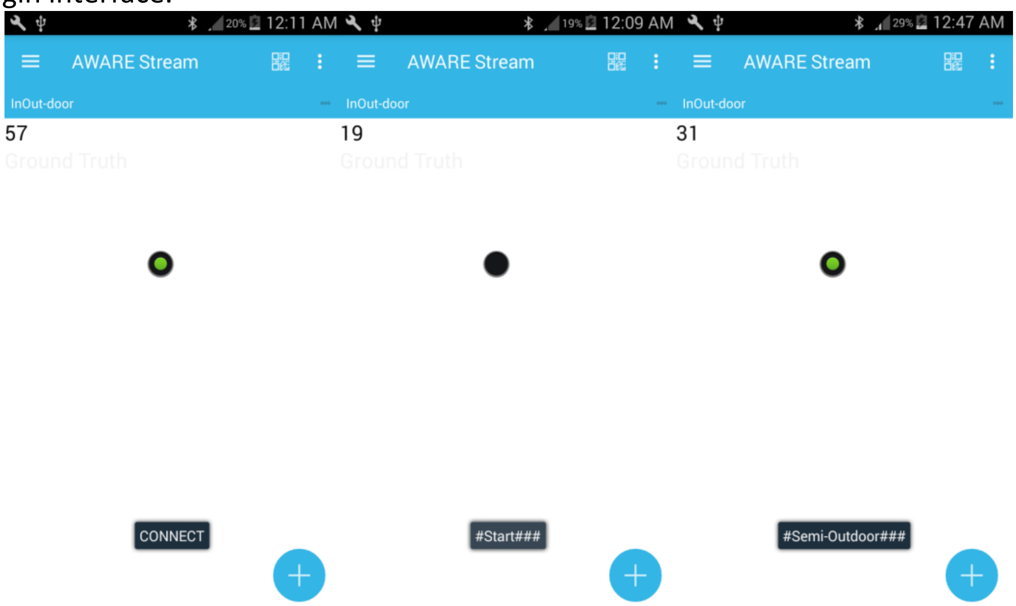


Figure 3. From left: the plugin has successfully received and accepted a connection from the remote labeller, labeller sent a start notification and labeller indicated that the device entered into a new environmental state (semi-outdoor).

3.2. Process and Environment

Data collection took place during 2015, in Oulu, Finland. A data collection session is called a trace, and during each trace the researcher collecting data visited all three environmental contexts at least once. Along the traces there were a rich variety of different surroundings such as a university campus, residential areas and rural areas with no nearby buildings. Each trace was collected using the mobile data collection toolkit described earlier. Figure 4 depicts examples of different areas where we collected data traces.

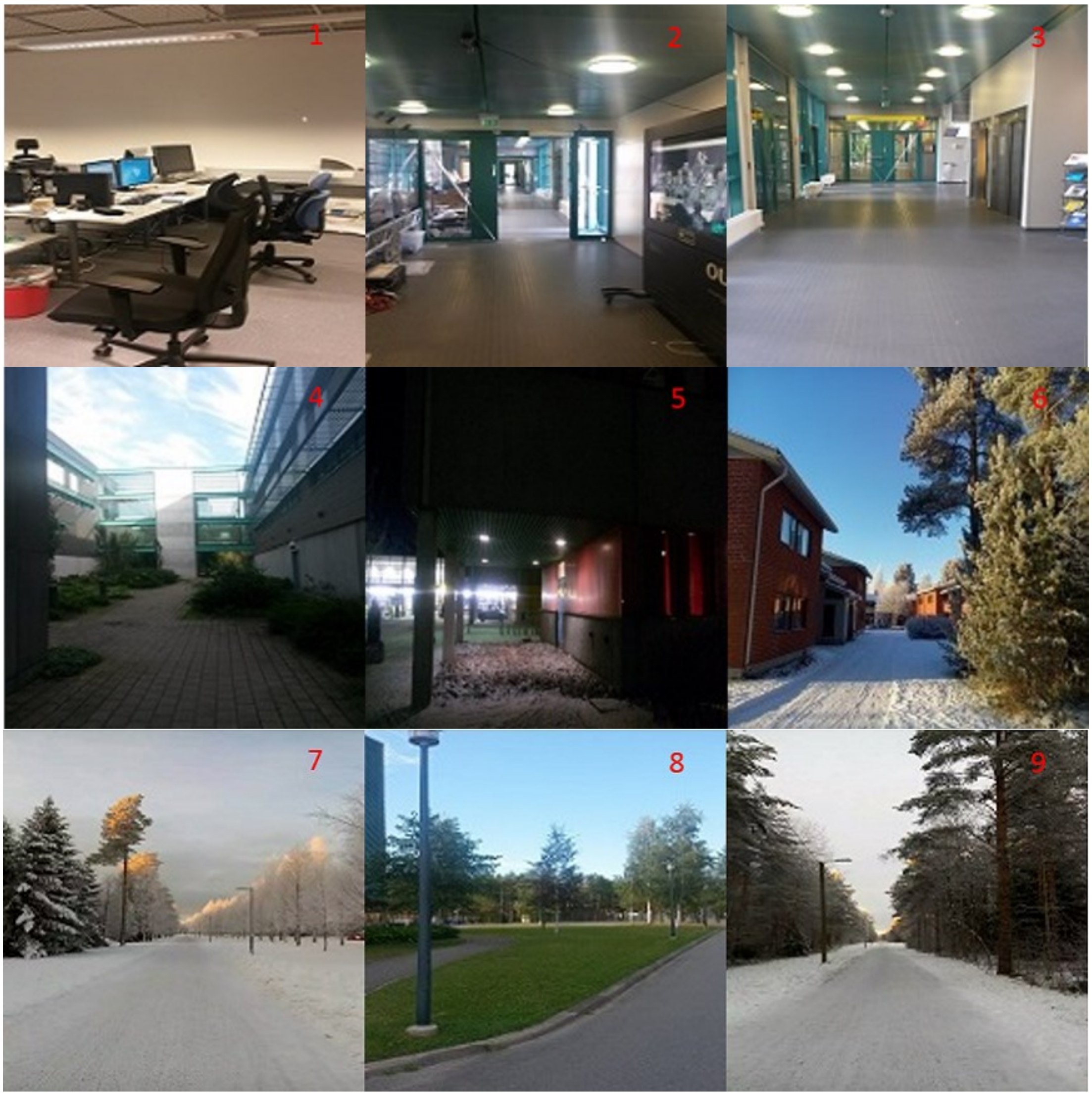


Figure 4. Different environments where we collected data for the analysis. Top row contains examples of indoor locations, middle row of semi-outdoor locations and bottom row of outdoor locations.

3.3. Results and Data Analysis

3.3.1. Collected data

We collected 10 data traces that altogether contain 52000 samples of magnetometer data, 7250 samples of accelerometer data, 30750 samples of light intensity data and 74 records of GSM cellular data. For the data analysis purposes of this thesis, we chose 2501 records of magnetometer data, 1300 records of accelerometer data, 1349 records of light intensity data and 25 records of GSM signal strength samples. These were chosen as they belong to specific traces that help best analyse changes in the environmental context.

3.3.2. Magnetometer analysis

In our analysis, we notice that magnetic field exhibits distinct patterns in different environments. In indoor area, for example, geomagnetic field is often disturbed by electric appliances and steel structures [50, 45]. Figure 5 plots the magnetic intensity during a data collection session where the researcher visited all environmental categories: indoor, semi-outdoor and outdoor. Specifically, in this scenario, a researcher holding a SAMSUNG GT19195 smart phone departs from a research laboratory to visit several semi-outdoor and outdoor spaces before finally arriving back to the laboratory. From the graph, we can see that in indoor areas the magnetic intensity varies more than that it does in semi-outdoor and outdoor areas. For instance, in the laboratory, from 90 seconds to 200 seconds, the intensity of magnetic field varies from about 0.15 to around 1.1 gauss; while from 220 seconds to 400 seconds, when in semi-outdoor and outdoor area, the intensity of magnetic field varies approximately from 0.6 to 0.85 gauss. The magnetic intensity variation in laboratory is approximately 0.95 gauss while that in semi-outdoor and outdoor is 0.25 gauss.

When the device is not moving and is indoor, the magnetic intensity does not show much variance. Further, in different places indoor, the magnetic field variation exhibits distinct patterns. For instance, in the laboratory, the magnetic field varies more dramatically than that does in normal buildings. The reason for this is that in laboratory there are several electric appliances that generate powerful electromagnetic fields that disturb the magnetic field.

Given the visual analysis based on the plotted magnetometer data, we find that it is possible to attempt distinguishing between the three environmental states simply by measuring and making a decision based on the magnetic field variation.

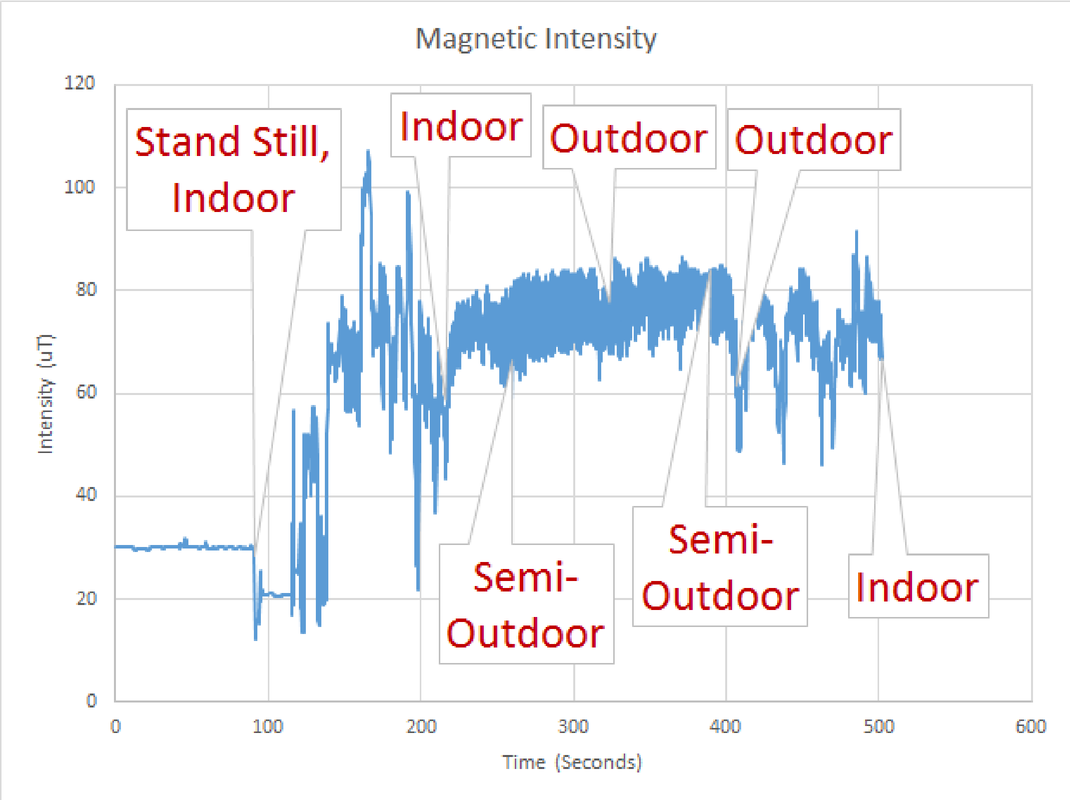
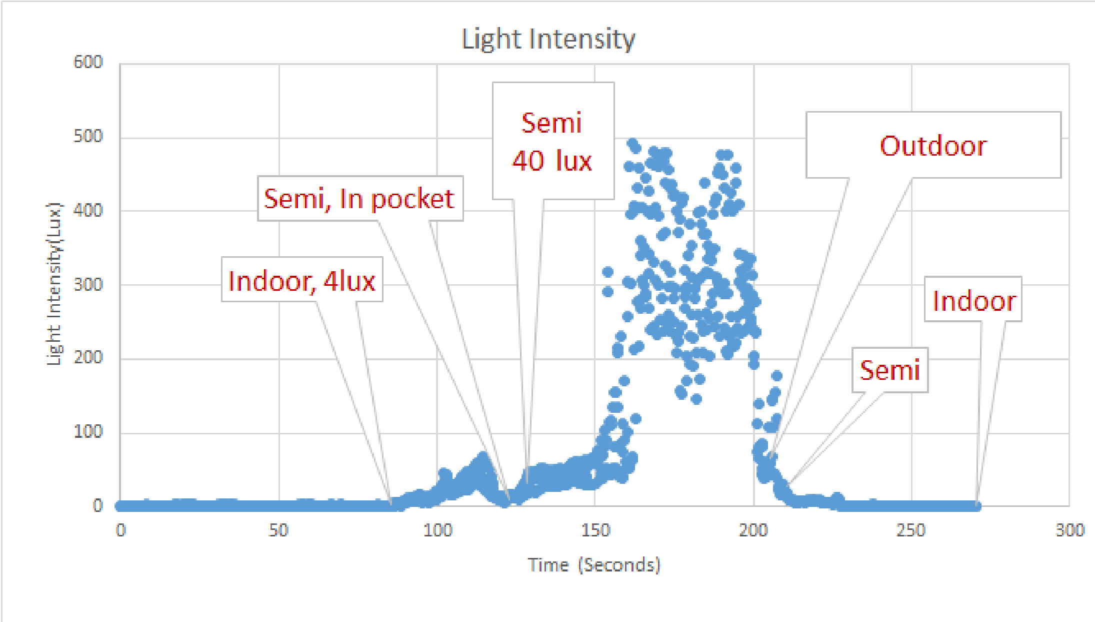


Figure 5. Magnetic field intensity plotted during one collected trace.

3.3.3. Light sensor analysis

Ambient light data is particularly challenging, as it is affected by several factors, such as weather, hour of day, seasons of year, artificial light sources, etc. First, we study outdoor light intensity in different situations (seasons, weather conditions, etc.). Figure 6 depicts light intensity during a data collection session. More specifically, during the depicted session the phone used for data collection was rotated towards different light sources. From 0 to 86 seconds, in indoor environment, the light intensity keeps below 10 Lux. From 86 seconds to 128 seconds, excluding the period when device is in pocket, in semi-outdoor environment, the light intensity is approximately from 10 to 70 Lux. In indoor environment the light intensity is above 100 Lux, ranging from 100 to 500 Lux. Only when in the interval from semi-outdoor to indoor area, the light intensity may appears to be lower than 100 Lux. Thus generally we conclude that light intensity outdoor is significantly higher than that in semi-outdoor environment where, however, the light intensity is still higher than in indoor environment. The reason behind this is because the natural sunlight intensity within visible spectrum is significantly higher than artificial light intensity[27].

Also it can be noticed that when the device is in pocket, light intensity decreases to near zero. Therefore, it is hard to infer indoor/outdoor/semi-outdoor environment when light intensity is very low, given that such situation may happen in any of these 3 environments.



Indoor Semi Outdoor Semi Indoor

Figure 6. Light intensity collected during one data collection session.

In addition, we show how light intensity varies between seasons. Figure 7 plots light intensity during winter and summer in Oulu, Finland (sunny summer in picture 5 in Figure 4 and cloudy winter in picture 9 in Figure 4). We notice that light intensity on a cloudy winter day is significantly weaker than during a sunny summer day in Oulu.

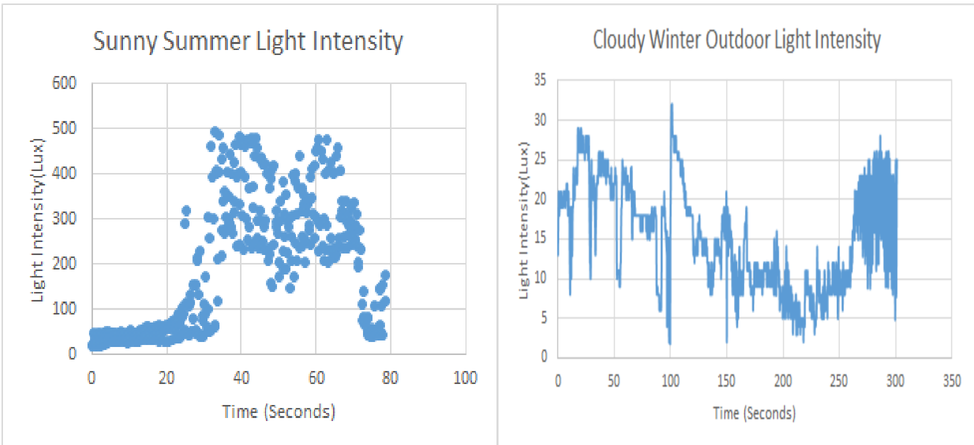


Figure 7. Outdoor Light Intensity in different environmental settings

To recap, light sensor seems highly useful in inferring indoor/semi-outdoor states, as in most cases outdoor light intensity is remarkably higher than indoor light intensity. However, if the intensity is low, it becomes hard to make any decisions of the context, as the low reading could be caused by the device being in a pocket, bag, etc. in any environment.

3.3.4. Cellular network analysis

To study the utility of GSM signal strength for environmental state detection we collected data from several different environments such as campus (semi-outdoor), corridors (semi-outdoor), playgrounds (outdoor), basketball courts (outdoor), offices (indoor) and laboratory premises(indoor). The absolute value of GSM strength in each environment does not have a general range and hence single GSM strength does not contribute to the detection of indoor/outdoor environment. Figure 8 illustrates this finding. When the environment abruptly (despite still being in the same indoor/outdoor context) changes, the cellular signal strength varies significantly as well [27]. The changes in signal strength in Figure 8 indicate such environment transitions. Still, the actual strength reading varies too much in different environments to be used as an indicator for indoor/outdoor contexts in our environment (Oulu, Finland).

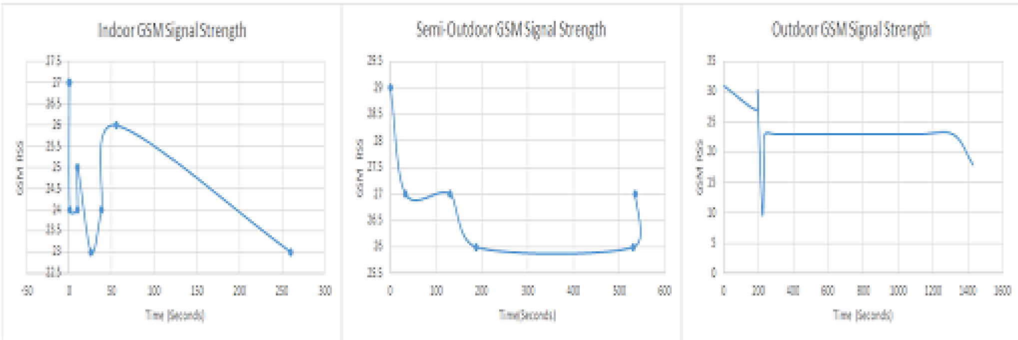


Figure 8. GSM signal strength in different environments.

3.3.5. Accelerometer analysis

Accelerometer measures the device's acceleration in 3 different physical dimensions (x, y, z axis). Here we note that Android's accelerometer includes the force of gravity. Figure 9 plots device acceleration (the joint acceleration generated from all the 3 dimensions) when the phone is idle and when the researcher is walking. Using accelerometer data alone it is impossible to infer environmental indoor/outdoor state, but it can, in theory, be used to make a more sophisticated guess when combined with other data sources. For example, when it is bad weather, it can be assumed that in general people do not stay still outside but are inside when not moving. Thus, accelerometer data may come handy in making a better overall decision about such states. Further, only when the phone is moving, magnetometer data becomes useable for environmental state inference, and thus accelerometer data can be used to define when magnetometer data should be used or not.

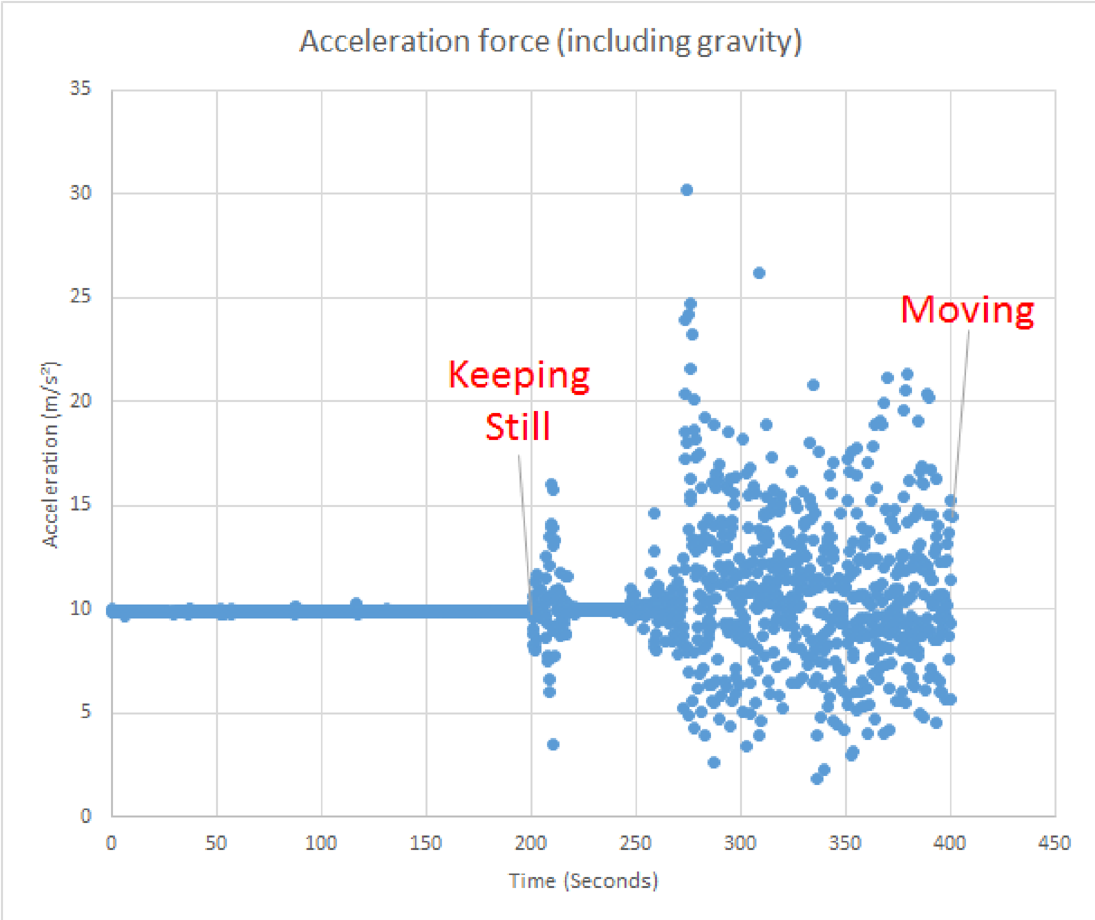


Figure 9: Acceleration value when first standing still and then walking.

3.4. Design Implications

The presented analysis informs the design of our own solution that is presented in the next chapter. First, we choose to use acceleration in combination with magnetometer data, as described earlier. When the phone is detected as moving, then we use variance of the magnetic field to infer indoor/outdoor state. For the accelerometer, we define so that if acceleration exceeds 11 m/s^2 or below 9 m/s^2 the phone is considered as moving.

Second, when light intensity is stronger than 100 Lux, we decide that the device is outdoor. The reason for this threshold (100 Lux) is that a clear majority of indoor and semi-outdoor data samples in our analysed light data are well below 100 Lux. Further, by observing the collected data we set 50 Lux as a threshold to differentiate semi-outdoor and indoor state. Anything below 50 Lux is considered as indoor or in pocket and anything above (or exactly) 50 Lux is considered to be semi-outdoor.

4. CONTEXTIO

Based on the design implications and other findings presented earlier we implemented a mobile application called ContextIO to enable real time indoor/outdoor classification on Android devices. ContextIO's key difference to related work is a different operational environment (northern city of Oulu, in Finland) and use of sensors that even as combined consume only very little energy.

4.1. System Overview

ContextIO uses only onboard sensors that most off-the-shelf smartphones have in 2016. The application consists of two conceptual main components that each use a combination of two input sensors. We denote the two components as *Magnetism Instrument* and *Light Instrument* (see Figure 10).

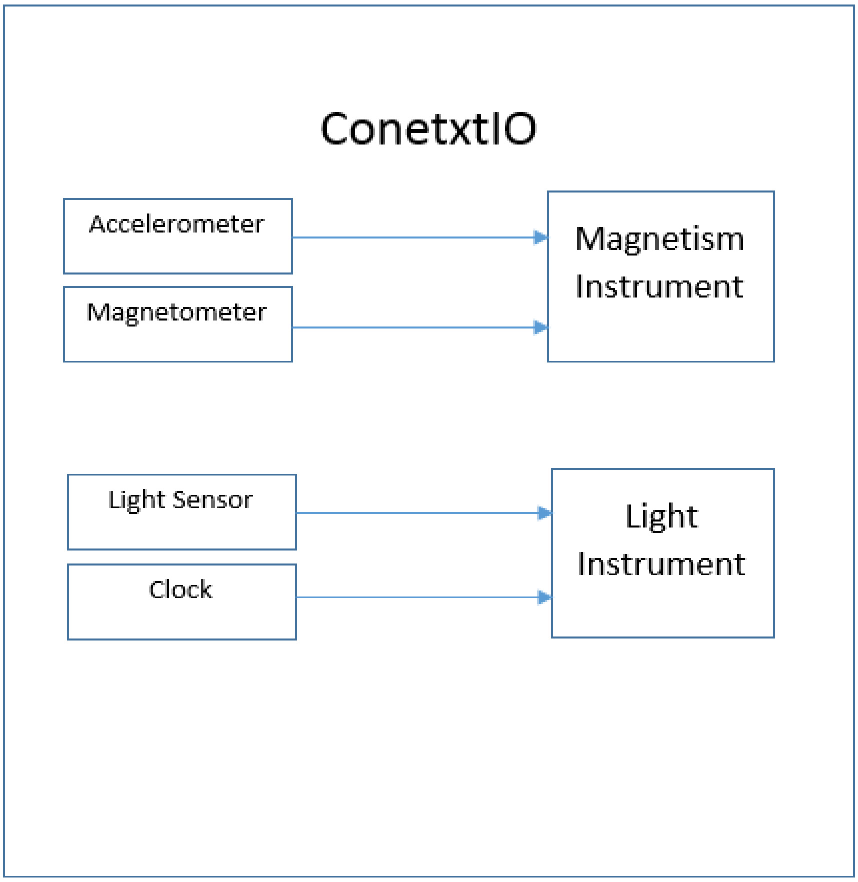


Figure 10. ContextIO conceptual diagram.

Magnetism Instrument uses a combination of accelerometer and magnetometer to detect the variation of magnetic intensity when mobile device is moving. In order to use magnetometer data, the accelerometer has to first detect the device as moving by using the thresholds defined earlier in section 3.4. *Light Instrument*, on the other hand, uses the onboard clock and light sensors to detect the time and measure the current light intensity. Clock is used, because only during daytime we can fully rely on light intensity as the state indicator. During night time, and especially during winter nights, light intensity is not very useable for state inference, as illustrated earlier in the data analysis. Figure 11 depicts the main user interface of ContextIO.

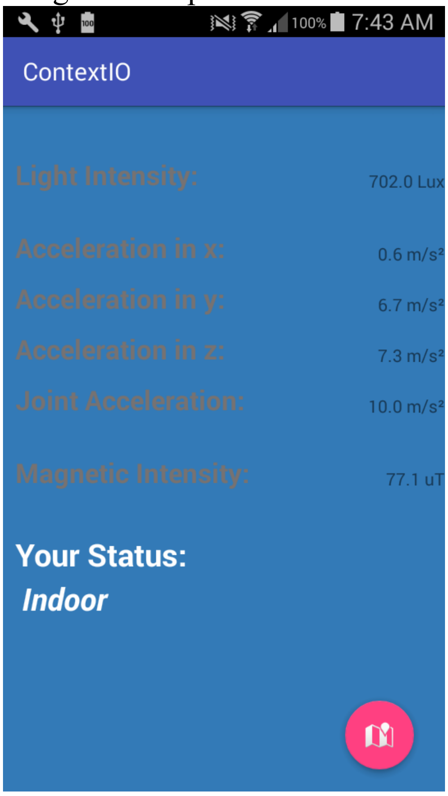


Figure 11. ContextIO user interface.

ContextIO uses magnetic field variance to detect between indoor and outdoor/semi-outdoor while light instruments are able to detect the outdoor and semi-outdoor. So *ContextIO* just uses magnetism to infer the state first, if the state is outdoor/semi-outdoor then using light instrument, it is possible to differentiate outdoor/semi-outdoor if current time is daytime. However, if the magnetism instrument has already detected that environment state is outdoor then there is no need to use light instrument anymore. The detection flow is shown in Figure 12.

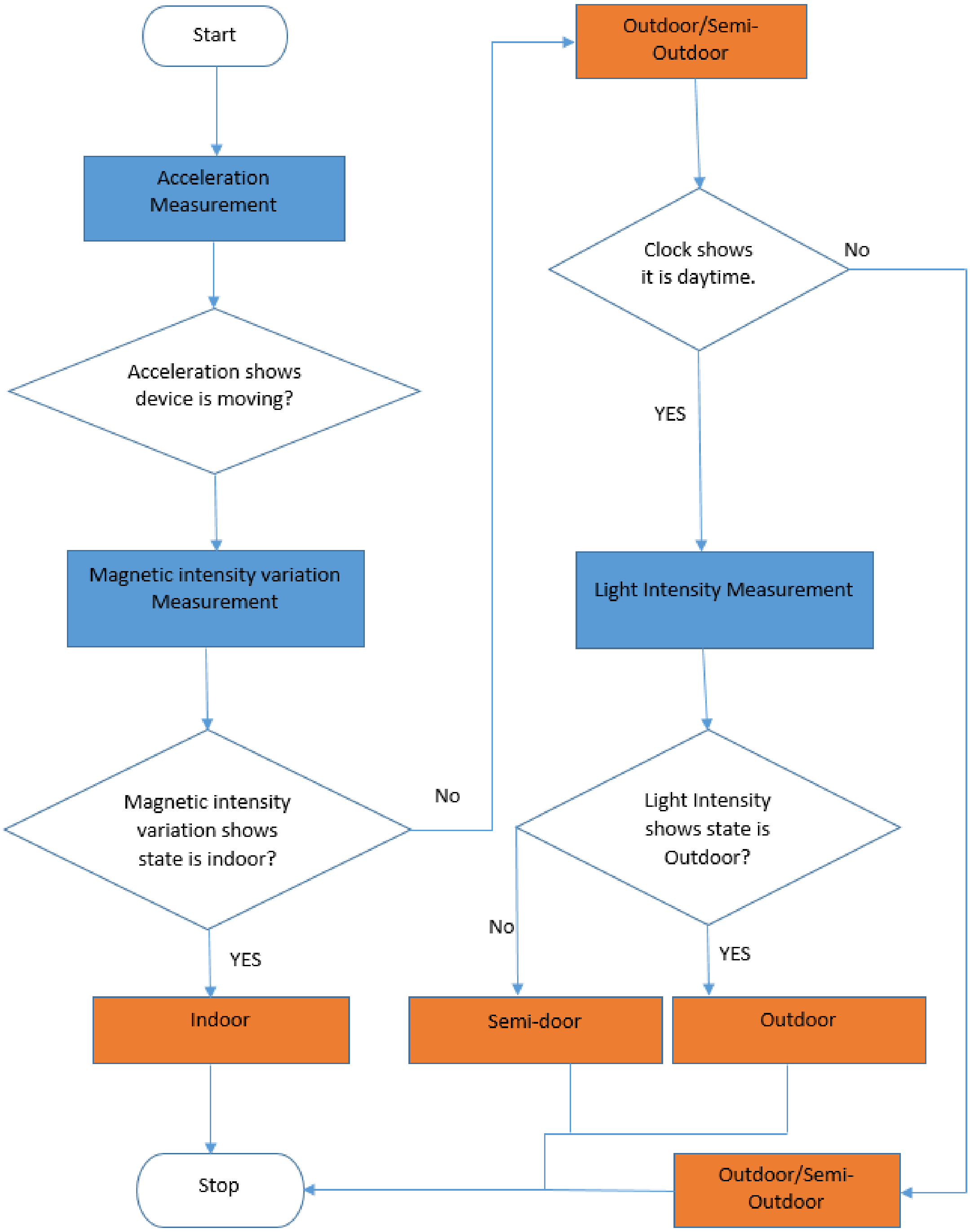


Figure 12. ContextIO state detection flow

4.2. ContextIO Accuracy

We have not systematically evaluated the final accuracy of ContextIO. However, we have completed several informative semi-formal test walks with the application in the same approximate context where the data was collected, and confirmed the accuracy to be approximately 60%. Give we are attempting to recognize between three states in a challenging environment, this is considered satisfactory at the moment. There exist several means to achieve higher accuracy (some are mentioned later in future work section), but we must also consider energy efficiency. Our solution, in theory, can achieve very low energy expenditure, given the sensors used. Further, as several related works have also noted, no solution generalizes flawlessly to other conditions. So, while we fully acknowledge the lack of formal accuracy evaluation, we also wish to highlight that such evaluations are almost without an exception bound to very specific conditions and therefore make it hard to compare between solutions in the first place.

5. DISCUSSION

5. 1. Energy Efficient and Unobtrusive Environmental State Detection

Several other techniques can be used to detect a device's environmental state. In related work, we discussed some of the most relevant ones to the thesis, and for example image processing and RFID-based solutions certainly compete well with our solution. However, such solutions require the user to do something, such as touch a tag or take a picture. We contribute a completely unobtrusive mechanism that can silently sense the state using onboard sensors and without requiring any extra effort from the user.

GPS and different types of fingerprint databases are other common solutions to state detection. However, for our purposes GPS is far from optimal for several reasons. First, the GPS receiver is very energy-intensive to use on mobile phones [32]. Second, the signal behaves very erratically in city centers and close to e.g. windows and open doors, making state inference challenging [31]. Fingerprint databases, on the other hand, require extensive data collection effort and a constant connection to the database to be kept open. Again, this is not very energy efficient and requires a lot of setup work to be performed.

5.2. Challenges and Future Work

Using mobile sensing to detect indoor/outdoor contexts is a topical research question with several acknowledged challenges. Similar to IODetector [27], our solutions suffers from the so called "adaptation problem", because we use manually hard-coded thresholds in the data for determining the context. Thus, the solution most likely will not generalise well to other specific locations and environments.

Our ongoing work considers machine learning to improve detection accuracy as well as to make the detection more sophisticated and adaptive to other environments. We have started exploring different machine learning models by using Weka [51]. Unlike in most previous works (e.g. [31, 46]), we aim to design a hybrid model where machine learning is used only for certain sensors and others, such as accelerometer, can still rely on manual thresholds. This simplifies the system as well as reduces energy expenditure. Our initial results are promising, and especially *Random Trees* seem to perform well with our approach.

Finally, we acknowledge limitations in our work. The data analysis is limited in scope, as 10 traces (and the analysed ones) certainly do not capture all the complexities involved in an authentic environment where people daily use their devices. However, we were able to analyse a sufficient set of realistic state transitions to clearly demonstrate how the sensors can be used in conjunction to infer states. Second, the analysis is not relying on acknowledge statistical methods, but is more a visual one. However, the differences in the plots are seemingly visible, and the thresholds defined clearly worked in determining the state when implemented in ContextIO.

5.3. Fulfilling the Objective and Answering Research Questions

The objective of this thesis was to create an Android application capable of detecting indoor/outdoor state of the device. ContextIO successfully does this, and its design is based on a data analysis concerning several sensors that can be in theory used for detecting such environmental states. To answer the first research question, we identified and tested several sensors available on modern Android phones that could be used to facilitate environmental state detection. Regards the second research question about the specific factors that influence the accuracy of the sensors, we identified the environment's effect on magnetometer data. Further, the effect of season and day of time must be taken into consideration when using light intensity for state detection.

6. CONCLUSION

This thesis focuses on modern mobile sensing means for detecting indoor/outdoor environmental states. First, we researched various techniques that have been used for such state detection, and chose the adequate ones to be explored more. More specifically, we chose to explore magnetometer, light sensor, cellular signal, and accelerometer, as all of them are relatively lightweight in terms of energy expenditure and most of them can be sensed using off-the-shelf smartphones.

Second, we developed a toolkit to collect data with the aforementioned sensors. The toolkit consists of 2 software components: 1) an AWARE plugin to facilitate data collection and ground truth labelling and 2) a remote labeller application to indicate ground truth states to the plugin.

We then analysed the collected data, and chose a set of sensors and manual thresholds that can together be used to infer indoor/outdoor states. Our findings show that when a smartphone is moving, magnetic intensity variation exhibits distinct patterns in indoor and outdoor and semi-outdoor states. Light sensors also have different range of variation in each of the 3 states. However, in our environment GSM data did not turn out very useful.

The, we developed a mobile solution, ContextIO, that relies on the findings of data analysis. Overall, ContextIO contributes to work on mobile environmental state detection by a lightweight, energy efficient mobile application and the accompanying data analysis presented in this thesis. Further, the approach of using a remote data labeller to indicate ground truth states is a novel and generalizable concept that can be used to e.g. collect natural light sensor data, as the collecting phone can be in a pocket or a bag. We hope our work to be inspirational in the context of mobile data sensing and especially for detecting different environmental states.

7. REFERENCES

[1] Wang Y, Lin J, Annavaram M, Jacobson QA, Hong J, Krishnamachari B, Sadeh N (2009) A framework of energy efficient mobile sensing for automatic user state recognition. 2009 Jun 22 In Proceedings of the 7th international conference on Mobile systems, applications, and services. Wroclaw, Poland, 179-192. DOI: http://dx.doi.org/10.1145/1555816.1555835

[2] Schilit B, Adams N, Want R (1994) Context-aware computing applications. 1994 In Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on 1994 Dec 8, Santa Cruz, CA, USA, 85-90. DOI: http://dx.doi.org/10.1109/MCSA.1994.512740

[3] Azizyan M, Constandache I, Roy Choudhury R (2009) SurroundSense: mobile phone localization via ambience fingerprinting. 2009 Sep 20 In Proceedings of the 15th annual international conference on Mobile computing and networking. Beijing, China , 261-272. DOI: http://dx.doi.org/10.1145/1614320.1614350

[4] Eriksson J, Girod L, Hull B, Newton R, Madden S, Balakrishnan H (2008) The pothole patrol: using a mobile sensor network for road surface monitoring. 2008 Jun 17 In Proceedings of the 6th international conference on Mobile systems, applications, and services. Breckenridge, CO, USA, 29-39. DOI: http://dx.doi.org/10.1145/1378600.1378605

[5] Eisenman SB, Miluzzo E, Lane ND, Peterson RA, Ahn GS, Campbell AT (2009) BikeNet: A mobile sensing system for cyclist experience mapping. ACM Transactions on Sensor Networks (TOSN). 6(1): 6. DOI: http://dx.doi.org/10.1145/1653760.1653766

[6] Ubicomp 2011 Workshop (2011) Mobile Sensing: Challenges, Opportunities and Future Directions. URL: http://research.microsoft.com/en-us/um/beijing/events/ms_ubicomp11/cfp.pdf. Accessed: 25, 10, 2016

[7] Hightower J, Borriello G (2001) Location sensing techniques. 2001 Jul 30 IEEE Computer. 34(8):57-66.

[8] Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT (2010) A survey of mobile phone sensing. IEEE Communications magazine 48(9):140-50. DOI: http://dx.doi.org/10.1109/MCOM.2010.5560598

[9] Cambridge English Dictionary URL: http://dictionary.cambridge.org/dictionary/english/context Accessed: 25, 10, 2016

[10] Schilit BN, Theimer MM (1994) Disseminating active map information to mobile hosts. IEEE network 8(5):22-32. DOI: http://dx.doi.org/10.1109/65.313011

[11] Chen Guanling & David Kotz (2000) A survey of context-aware mobile computing research. Technical Report TR2000-381. Dept. of Computer Science, Dartmouth College.

[12] Abowd GD, Dey AK, Brown PJ, Davies N, Smith M, Steggles P (1999) Towards a better understanding of context and context-awareness. In International Symposium on Handheld and Ubiquitous Computing 1999 Sep 27, Karlsruhe, Germany, 304-307. DOI: http://dx.doi.org/10.1007/3-540-48157-5_29

[13] Ballagas R, Borchers J, Rohs M, Sheridan JG (2006) The smart phone: a ubiquitous input device. IEEE Pervasive Computing. 2006 Jan; 5(1):70-7. DOI: http://dx.doi.org/10.1109/MPRV.2006.18

[14] Rochelle Garner (2016) Nearly 9 out of 10 smartphones shipped run on Android. URL: https://www.cnet.com/news/android-smartphone-shipments-88-percent-marketshare-strategy-analytics/ Accessed: 28. 10. 2016

[15] Android development API guides. URL: https://developer.android.com/guide/topics/sensors/sensors_overview.html Accessed: 28. 10. 2016

[16] Zhen C, Qiang G (2013) Mobile Sensor Data Collecting System Based on Smart Phone. 2013 Dec 5 In Joint International Conference on Pervasive Computing and the Networked World. Vina del Mar, Chile, 8-14. DOI: http://dx.doi.org/10.1007/978-3-319-09265-2_2

[17] WIKIPEDIA: internet of things. URL: https://en.wikipedia.org/wiki/Internet_of_things Accessed: 1. 11. 2016

[18] Hammad Tariq (2016) Collecting smartphone sensor data via PubNub messaging. URL: https://evothings.com/collecting-phone-sensor-data-on-pubnub/ Accessed: 1. 11. 2016

[19] Yi WJ, Jia W, Saniie J (2012) Mobile sensor data collector using Android smartphone. In 2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS) 2012 Aug 5. Boise, Idaho, USA, 956-959. DOI: http://dx.doi.org/10.1109/MWSCAS.2012.6292180

[20] Lane ND, Chon Y, Zhou L, Zhang Y, Li F, Kim D, Ding G, Zhao F, Cha H (2013) Piggyback CrowdSensing (PCS): energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities. 2013 Nov 11 In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems. Roma, Italy, No.7. DOI: http://dx.doi.org/10.1145/2517351.2517372

[21] Priyantha B, Lymberopoulos D, Liu J (2011) Littlerock: Enabling energy-efficient continuous sensing on mobile phones. IEEE Pervasive Computing 10(2):12-5. DOI: http://dx.doi.org/10.1109/MPRV.2011.28

[22] Denzil Ferreira (2016) AWARE's Introduction URL: http://www.awareframework.com/what-is-aware/ Accessed: 20. 10. 2016

[23] Hosio S, Ferreira D, Goncalves J, van Berkel N, Luo C, Ahmed M, Flores H & Kostakos V (2016) Monetary Assessment of Battery Life on Smartphones. Proc. 2016 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI). San Jose, CA, USA, 1869-1880. DOI: http://dx.doi.org/10.1145/2858036.2858285

[24] van Berkel N, Luo C, Anagnostopoulos T, Ferreira D, Goncalves J, Hosio S & Kostakos V (2016) A Systematic Assessment of Smartphone Usage Gaps. Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI). San Jose, CA, USA, 4711-4721. DOI: http://dx.doi.org/10.1145/2858036.2858348

[25] Froehlich J, Chen MY, Consolvo S, Harrison B, Landay JA (2007) MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. 2007 Jun 13 In Proceedings of the 5th international conference on Mobile systems, applications and services. PR, USA, 57-70. DOI: http://dx.doi.org/10.1145/1247660.1247670

[26] Aharony N, Pan W, Ip C, Khayal I, Pentland A (2011) The social fMRI: measuring, understanding, and designing social mechanisms in the real world. 2011 Sep 17 In Proceedings of the 13th international conference on Ubiquitous computing. Beijing, China, 445-454. DOI: http://dx.doi.org/10.1145/2030112.2030171

[27] Zhou P, Zheng Y, Li Z, Li M, & Shen G (2012, November) IODetector: a generic service for indoor outdoor detection. 2012 In Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems. Toronto, Canada, 113-126. DOI: https://dx.doi.org/10.1145/2426656.2426668

[28] Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT (2010) A survey of mobile phone sensing. IEEE Communications magazine 48(9):140-50. DOI: http://dx.doi.org/10.1109/MCOM.2010.5560598

[29] Qin C, Bao X, Roy Choudhury R, Nelakuditi S (2011) Tagsense: a smartphone-based approach to automatic image tagging. 2011 In Proceedings of the 9th international conference on Mobile systems, applications, and services Bethesda, MD, USA, 1-14. DOI: http://dx.doi.org/10.1145/1999995.1999997

[30] Szummer M, Picard RW (1998) Indoor-outdoor image classification. Jan 3, 1998 InContent-Based Access of Image and Video Database. Bombay, India. 42-51. DOI: http://dx.doi.org/10.1604/9780818683299

[31] Radu V, Katsikouli P, Sarkar R, Marina MK (2014) Poster: am i indoor or outdoor?. 2014 Sep 7 In Proceedings of the 20th annual international conference on Mobile computing and networking. New York, NY, USA, 401-404. DOI: http://dx.doi.org/10.1145/2639108.2642916

[32] Ben Abdesslem, Fehmi, Andrew Phillips, and Tristan Henderson (2009) Less is more: energy-efficient mobile sensing with senseless. 2009 August 16 - 21 Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds. Barcelona, Spain, 61-62. DOI: http://dx.doi.org/ 10.1145/1592606.1592621.

[33] Chon HD, Jun S, Jung H, An SW (2004) Using RFID for accurate positioning. Journal of Global Positioning Systems. 1(08): 32-39. DOI: https://dx.doi.org/10.5081/jgps.3.1.32

[34] Chon HD, Jun S, Jung H, An SW (2004) Using RFID for accurate positioning. Journal of Global Positioning Systems. 1(08): 32-39. DOI: https://dx.doi.org/10.5081/jgps.3.1.32

[35] Want R (2006) An introduction to RFID technology. IEEE Pervasive Computing. 2006 Jan; 5(1): 25-33. DOI: https://dx.doi.org/10.1109/MPRV.2006.2

[36] Barnes J, Rizos C, Wang J, Small D, Voigt G, Gambale N (2002 Dec) High precision indoor and outdoor positioning using LocataNet. Journal of Global Positioning Systems (2003) 2(2): 73-82. DOI: https://dx.doi.org/10.5081/jgps.2.2.73

[37] Randell C, Muller H (2001) Low cost indoor positioning system. 2001 Sep 30 In International Conference on Ubiquitous Computing. Atlanta Georgia, USA, 42-48. DOI: https://dx.doi.org/10.1007/3-540-45427-6_5

[38] Yang Z, Wu C, Liu Y (2012) Locating in fingerprint space: wireless indoor localization with little human intervention. 2012 In Proceedings of the 18th annual international conference on Mobile computing and networking. Istanbul, Turkey, 269-280. DOI: https://dx.doi.org/10.1145/2348543.2348578

[39] Zhang Z, Zhou X, Zhang W, Zhang Y, Wang G, Zhao BY, Zheng H (2011) I am the antenna: accurate outdoor AP location using smartphones. 2011 In Proceedings of the 17th annual international conference on Mobile computing and networking. Las Vegas, NV, USA, 109-120. DOI: https://dx.doi.org/10.1145/2030613.2030626

[40] Kaemarungsi K, Krishnamurthy P (2004) Properties of indoor received signal strength for WLAN location fingerprinting. 2004 Aug 22 In Mobile and Ubiquitous Systems: Networking and Services, 2004.

MOBIQUITOUS 2004. The First Annual International Conference. Boston, Massachussets, USA,14-23. DOI: http://doi.ieeecomputersociety.org/10.1109/MOBIQ.2004.1331706

[41] Gani MO, OBrien C, Ahamed SI, Smith RO (2013) RSSI based indoor localization for smartphone using fixed and mobile wireless node. 2013 Jul 22 In Computer Software and Applications Conference (COMPSAC), 2013 IEEE 37th Annual. Kyoto, Japan, 110-117. DOI: https://dx.doi.org/10.1109/COMPSAC.2013.18

[42] Leu JS, Yu MC, Tzeng HJ (2015) Improving indoor positioning precision by using received signal strength fingerprint and footprint based on weighted ambient Wi-Fi signals. Computer Networks. 91:329-40. DOI: http://dx.doi.org/10.1016/j.comnet.2015.08.032

[43] Wang F, Huang Z, Yu H, Tian X, Wang X, Huang J (2013) EESM-based fingerprint algorithm for Wi-Fi indoor positioning system. In 2013 IEEE/CIC International Conference on Communications in China (ICCC) 2013 Aug 12. Xi'an, China, 674-679. DOI: http://dx.doi.org/10.1109/ICCChina.2013.6671197

[44] Roberts G, Anvari K & inventors (1992) Novatel Communications, Ltd., assignee. Cellular radio-telephone receiver employing improved technique for generating an indication of received signal strength. United States patent US 5,142,695.

[45] Chung J, Donahoe M, Schmandt C, Kim IJ, Razavai P, Wiseman M (2011) Indoor location sensing using geo-magnetism. 2011 Jun 28 In Proceedings of the 9th international conference on Mobile systems, applications, and services. Bethesda, MD, USA, 141-154. DOI: https://dx.doi.org/10.1145/1999995.2000010

[46] Radu V, Katsikouli P, Sarkar R, Marina MK (2014) A semi-supervised learning approach for robust indoor-outdoor detection with smartphones. 2014 Nov 3 In Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems. Memphis, TN, USA, 280-294. DOI: https://dx.doi.org/10.1145/2668332.2668347

[47] Kroll N, Jäger M, Süß S. Context-aware Indoor-Outdoor Detection for Seamless Smartphone Positioning.

[48] Hansen R, Wind R, Jensen CS, Thomsen B (2009) Seamless indoor/outdoor positioning handover for location-based services in streamspin. In 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware 2009 May 18. Taipei, Taiwan, 267-272. DOI: http://dx.doi.org/10.1109/MDM.2009.39

[49] Ferreira D (2013) Aware: A Mobile Context Instrumentation Middleware to Collaboratively Understand Human Behavior." PhD Thesis. University of Oulu, Department of Computer Science and Engineering.

[50] Rowe A, Gupta V, Rajkumar RR (2009) Low-power clock synchronization using electromagnetic energy radiating from ac power lines. 2009 Nov 4 In Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems. New York, USA, 211-224. DOI: http://dx.doi.org/10.1145/1644038.1644060

[51] Machine learning Group at the University of Waikato (2016) WEKA Introduction http://www.cs.waikato.ac.nz/ml/weka/ Accessed: 20. 11. 2016