Translating mobile sensor data
Into meaningful context descriptions

Magdalena Mucko
s060665

April 30, 2010
Informatics and Mathematical Modelling, IMM
Technical University of Denmark, DTU
Supervised by Jakob Eg Larsen and Michael Kai Petersen
Abstract

Sensor enabled devices (in case of this thesis - mobile phones) can capture and process individual’s context information and transmit this data to remote locations without any human intervention. Everyone who has ever tried using the Internet for some time is aware of the social networking phenomenon. Such sites as Facebook, LinkedIn, and MySpace, as well as content-sharing services that also offer social networking functionality (including YouTube, Flickr) have captured the attention of millions of users.

This thesis approaches the concept of merging users’ context data with their social networking life. This work specifically focuses on utilization of the context information to deliver the most up to date knowledge about the user’s condition. That statement creates a foundation for development of the “ContextBox” application, which employs a small group of selected sensors, to deliver a concise view on the user’s current status.

Implementation of a prototype application and carrying out a user study allowed for a preliminary analysis of potential usability of selected solution and its research value.
Preface

This thesis was prepared at the Department of Informatics and Mathematical Modeling (IMM), at the Technical University of Denmark (DTU) in partial fulfillment of the requirements for acquiring the degree of Master of Science in Telecommunication.

The thesis supervisors are Jakob Eg Larsen and Michael Kai Petersen, Department of Informatics and Mathematical Modeling, Technical University of Denmark.
# Table of Contents

Abstract ............................................................................................................................................. i  
Preface............................................................................................................................................. iii  
Table of Contents ............................................................................................................................. v  
List of Figures .................................................................................................................................. vii  
List of Tables .................................................................................................................................... ix  
List of Code snippets ....................................................................................................................... ix  
1 Introduction .................................................................................................................................. 1  
1.1 Background / Motivation ...................................................................................................... 1  
1.2 Problem Definition .............................................................................................................. 1  
1.3 Thesis Structure .................................................................................................................... 2  
2 Analysis ........................................................................................................................................ 5  
2.1 Predicting human behavior ................................................................................................. 5  
2.2 Context-awareness .............................................................................................................. 6  
2.2.1 Context .......................................................................................................................... 7  
2.2.2 The role of sensors ....................................................................................................... 9  
2.3 The Social Networking Sites phenomenon ........................................................................ 13  
2.3.1 The most popular Social Network Site ..................................................................... 13  
2.3.2 Bringing context-awareness into the Social Networking Sites .................................. 16  
2.4 Privacy issues ...................................................................................................................... 18  
2.4.1 Private data presented online ...................................................................................... 19  
2.4.2 Threats .......................................................................................................................... 19  
2.5 Summary .............................................................................................................................. 21  
3 Design ......................................................................................................................................... 23  
3.1 General idea .......................................................................................................................... 23  
3.2 Assumptions ........................................................................................................................ 23  
3.3 ContextBox application ........................................................................................................ 24  
3.3.1 Collecting data from phone’s sensors ......................................................................... 25  
3.3.2 Hosting the application and storing the data ............................................................... 28  
3.3.3 Social Networking platform as a context presentation layer ...................................... 32  
3.4 Communication flow .......................................................................................................... 33
4 Implementation .......................................................................................................................... 35
  4.1 ContextBox components ........................................................................................................ 35
    4.1.1 Mobile Context Toolbox ................................................................................................. 36
    4.1.2 Google App Engine service ............................................................................................ 37
  4.2 Facebook platform .................................................................................................................... 44
    4.2.1 Graphical presentation of the context information ....................................................... 45
    4.2.2 Privacy settings ............................................................................................................. 46
5 Evaluation ..................................................................................................................................... 49
  5.1 ContextBox evaluation ............................................................................................................. 49
  5.2 Survey .................................................................................................................................. 51
    5.2.1 General satisfaction level of Facebook ............................................................................ 51
    5.2.2 Facebook usage scenarios ............................................................................................. 52
    5.2.3 Privacy specific settings of Facebook ............................................................................. 53
    5.2.4 Common Facebook activities ......................................................................................... 54
    5.2.5 Sharing context information on Facebook ...................................................................... 54
    5.2.6 Summary ...................................................................................................................... 57
6 Discussion and Conclusions .......................................................................................................... 59
References ....................................................................................................................................... 63
Appendix A ..................................................................................................................................... 67
Appendix B ..................................................................................................................................... 69
List of Figures

Figure 1: Levels of context ............................................................................................................... 8
Figure 2: Client localization using BTS stations .............................................................................. 10
Figure 3: ContextBox system architecture ..................................................................................... 24
Figure 4: Information level Use Case Diagram ............................................................................... 26
Figure 5: Context Toolbox Architecture ........................................................................................ 27
Figure 6: Knowledge (blue color) and wisdom (green color) levels Use Case Diagram ................. 29
Figure 7: Typical user profile site .................................................................................................... 32
Figure 8: Information flow sequence diagram ............................................................................... 34
Figure 9: ContextBox components ................................................................................................. 35
Figure 10: User’s last 10 known GSM locations .............................................................................. 43
Figure 11: GSM to WiFi timestamp mapping and lookup time window expansion ....................... 44
Figure 12: Facebook application architecture ................................................................................ 44
Figure 13: Battery state icons ......................................................................................................... 45
Figure 14: Profile icons ................................................................................................................... 45
Figure 15: People around/Alone icons ........................................................................................... 45
Figure 16: Location icon ................................................................................................................. 46
Figure 17: Life designing with Firebug ........................................................................................... 46
Figure 18: Processed context information (narrow ContextBox) on the user’s profile page ........... 49
Figure 19: Satisfaction statistics (“How satisfied are you with Facebook, overall?”) .................... 52
Figure 20: Average time per day spend on Facebook .................................................................... 52
Figure 21: Facebook as an everyday activity .................................................................................. 52
Figure 22: Privacy settings .............................................................................................................. 53
Figure 23: Providing email address and phone number on Facebook ........................................... 53
Figure 24: Willingness to share context information on Facebook ............................................... 54
Figure 25: The way of collecting sensor data ................................................................................. 54
Figure 26: Type of information to be presented on Facebook ...................................................... 55
Figure 27: Willingness to share location information .................................................................... 55
Figure 28: Willingness to share location history ......................................................................... 55
Figure 29: Willingness to share activity information ..................................................................... 56
Figure 30: Willingness to share "people around" information ...................................................... 56
Figure 31: Willingness to share calendar events information ....................................................... 56
Figure 32: Willingness to see friends’ context information ......................................................... 56
Figure 33: Type of information to be seen on friends’ Facebook profiles ..................................... 57
List of Tables

Table 1: Summary of provided APIs ............................................................................................... 14
Table 2: Facebook and MySpace comparison ................................................................................ 15
Table 3: Get sensor data use case .................................................................................................. 25
Table 4: Infer meaningful places use case ...................................................................................... 30
Table 5: Infer if user is busy use case ............................................................................................. 30
Table 6: Infer activity use case ....................................................................................................... 31
Table 7: Infer relationship type use case ........................................................................................ 31
Table 8: Present information use case ........................................................................................... 32

List of Code snippets

Code snippets 1: Bluetooth model ................................................................................................. 38
Code snippets 2: Sample log data JSON notation .......................................................................... 39
Code snippets 3: Sample data submitted to Google geolocation service ........................................ 41
Code snippets 4: Sample Google geolocation response ................................................................. 42
1 Introduction

1.1 Background / Motivation

Information about the human environment offers new opportunities in terms of time-aware, location-aware, device-aware and personalized applications. Such systems constantly need to monitor the user’s surroundings, and here comes the role of sensors. Sensor devices are being embedded in all sorts of objects including vehicles, furniture but also animal and human bodies through health monitors and tagging techniques. Sensors are also becoming more common in mobile devices. Within the recent few years, mobile phones transformed from simple communication devices into advanced mobile computing platforms. Many of the mobile devices are already equipped with a camera, a GPS receiver, an accelerometer, a light sensor, etc. Moreover, year by year, the Internet transformed from just an endless source of information, into a communication tool. Most of the people, especially the young one, cannot imagine their lives without continuous communication with their friends. They call each other, send SMS, share photos and other personal information on the existing social networks.

Social Networking Sites are competing with older communication modalities, such as: voice (phone) calls, SMS messages, e-mails, and chat sessions. As our online and offline lives become inextricably linked, more and more people are sharing the details of their daily lives for business and pleasure. An example of social networks can be Facebook and MySpace that, allow users to create personalized profiles that are linked to their friends.

A mobile phone accompanies us almost everywhere and at the same time its sensors are able to collect interesting information about users’ environment. The possibilities of the phone can be connected with someone's favorite social network, where dozens of friends keep tracking each others’ profiles, and at the same time could help avoiding unnecessary (and sometimes annoying) questions: “Where are you?”, “What are you doing”, “Can we talk right now?”, etc.

The integration of sensor networks with social networks leads to applications that can sense the context of a user in much better ways and thus provides more personalized and detailed solutions.

1.2 Problem Definition

Social networks became a phenomenon of the present-days. Almost every one (regardless of age) has an account on at least one of the Internet society sites.
People are sharing their interests, photos, thoughts. They can sit in front of their computers and watch other people's lives. With the closest of the friends they are in touch not only through the social network but also through the mobile phone.

The most of the Social Network Sites usually offer the same basic functionalities: friends' listings, person surfing, private messaging, discussion forums or communities, events management, blogging, commenting, and media uploading. Those features can be enriched by adding the user's context, which represents his current status.

For someone who is sitting at home alone, the only way to ask a friend what he or she is doing and if he/she is in the neighborhood is to make a phone call. What if the friend is not picking up the phone or picks it up and whispers that he/she cannot talk right now. It can be a bit embarrassing to know that you bother someone without having an important reason.

In this thesis a concept of feeding the context information into the social network is presented. Mobile phone is used as a “medium” between the user and one of the social networks. All of the sensors which are embedded in the phone and some of its functionality enable gathering information about the user's status. The goal is to give meaning to that low-level context data that is legible and easy to understand. Then authorized friends from the social network would be able to see on the user's profile the interpreted context information that was sent by the mobile phone.

The main goal is to develop and present a prototype of a mobile phone application that cooperates with a chosen social network. Application is expected not to interfere too much with the user.

The thesis would include an overview of the existing solutions which would help to investigate, which of the context information is the most useful to the user. The processed data will be send to the social network site and will be displayed in a user friendly way.

1.3 Thesis Structure

The report is divided into chapters which represent following steps in the project realization. At the beginning an analysis of the topic has been conducted. Available sensors in the mobile phones along with their utilization and limitations have been shown. The definition of a user’s context has been presented. Additionally, enormous popularity of social network sites, as well as its privacy problems has been described. The next chapter – Design, explains the general concept of the project. The application design has been introduced. The next section is an Implementation chapter where a detailed description of the selected methods has been shown.
The Evaluation chapter illustrates the functionality of the application. Moreover, a survey that examines the general users’ opinion on Facebook and possibility of sharing context has been presented. Finally, the Discussion and Conclusions chapter is dedicated to observations made while working on this thesis.
2 Analysis

In this chapter a topic of patterns observed in human’s life will be brought up. Then the context-awareness of applications will be defined. Additionally, mobile phones’ sensors along with its sample utilization and chosen existing sensor based services will be presented. A short overview of selected Social Network Sites (SNS) will follow. The networks will be analyzed in terms of relations between the members. Finally the topic of privacy issues will be explained.

2.1 Predicting human behavior

People think that they are active, unpredictable creatures. It’s almost not possible to know precisely where everyone is all the time, however mobile phones are able to provide a good approximation. These small devices are following people almost all the time and are collecting data, that average user is even not aware of.

Cell phone companies store records of customers’ whereabouts. This information is based on location of the cell towers customers’ phones connect during making calls. A new report that tracked mobile phone users’ locations proved that humans’ everyday whereabouts can be predicted up to 93% of the time (1). The research shows that people are following the same travel routines and they are visiting the same places on a regular basis. A history of the visited places, a time spent at an area of each cell tower, an order in which phone’s user is changing his/her location, all this data helps in carrying out the research on finding the patterns in humans’ life. Additionally it doesn't matter if phones’ users stick to the same small area or travel a lot. Everyone is almost equally predictable, regardless of the size of their traveled region.

By combining low-level sensor data (for example GPS\(^1\) coordinates) with general commonsense knowledge of real-world constraints, it is possible to build a predictive model of human behavior (2). The commonsense knowledge can be considered as some of the real-world limitations, for example, all cars and busses can only move on the streets and trains on the railways, passengers can enter busses and trains only at determined stops, etc. Additionally, information such as the day of the week or time may be really significant for the prediction of life patterns.

Watching the history of the data collected by sensors (e.g. build in the mobile phone) enable to prove that ordinary person’s life is full of patterns. World is forcing humans to live in a predetermined manner. During the week (from Monday to Friday) almost everyone in the morning is travelling form home to work or school, and for most of the year people spend nights

\(^{1}\) Global Positioning System
at their homes. Discovering the human’s behavior model is more difficult when the observed person is in a new location (never been in this place before).

Location is not the only information in this topic that has been studied. Humans’ activity sequences have been also a part of the research. Considering that people have 24-hour periodic daily routines, long-range dependencies between activities are more significant for finding the prediction scheme (3). It seems to be logical that if someone has been working till late hours at night, he probably wake up late in the morning as well as if he takes a morning bus to work, he certainly came back home in the evening also by bus. An example represents a project (3), during which sensor data form 94 mobile phones (call logs, Bluetooth devices in proximity, phone status, cell tower ids, and application usage) have been collected and analyzed. The research proves that the daily activities are significantly related (e.g. what user does at 7-8 pm is linked with what he does in the morning and what he will do at 6 am depends on what she was doing at 2 am).

Apart from location and activity, a topic that is worth mentioning, is predicting the user’s state and the possibility of interrupting in something important (4). An individual looking for someone else’s attention can quickly notice if he/she is disturbing or not. Knowledge of the current situation enables the possibility of waiting for the appropriate moment. This behavior is considered as natural, socially appropriate and polite. Unfortunately phones and other electronic devices cannot naturally act in the same way. This comes down to forcing the device to be silent or simply turning it off. To make life easier for the user the mobile phones could automatically recognize specific places or situations and change the profile to the suitable one. Additionally, “potential intruders” could be also informed that it’s not the best time for calling.

2.2 Context-awareness

The modern mobile phones (5)(6)(7) are equipped with sensors such us GPS, Wi-Fi, cellular localization, Bluetooth, accelerometer, digital compass and etc.

As those small devices have become an integral part of people daily lives, it means that build into sensors are with the user at all time, collecting the information and allowing the system to be “aware” of the user’s context. The definition of “context” is presented in the next section.
2.2.1 Context

The “context-awareness” can be utilized to improve the social communication. The key to offer personalized and relevant service is to understand the “context” definition (8):

“Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.”

Simplifying, the context is any piece of information about people’s environment and actions in time. Below there are listed some of the most significant components of a user’s context (9):

- Personal information – e.g. interests
- Social information – e.g. friends, relationship, health
- Schedule – e.g. calendar, meetings, to-do list
- Present status – e.g. location, activity, not available
- Presence and availability – e.g. busy, not available
- Device capabilities – e.g. battery state
- Subscriptions
- Privacy and access control

The data, that seems to be the easiest to capture by sensors and convert into meaningful content, are user’s routines (10) (temporal regularities in his life). As it was already mentioned in chapter 2.1, many activities are interrelated (e.g. being at work is usually followed by going home) or form patterns (e.g. gym on Tuesdays, pub on Saturday). When the computing system has the knowledge about a user’s context it can predict his/her current situation (by deducing and combining facts).

The Figure 1 presents the level of context (8). The bottom layer represents raw sensor data. Mobile phone can collect miscellaneous information, which separately very often means nothing. Knowing only a time of the day, only MAC address of a nearby device or just coordinates of the place it is hard to deduce anything. By increasing the quantity of information and focusing on the relation between it, it is possible to step into the 2nd level of the context. After studying the gathered data, patterns in human life start to appear. By monitoring location in connection with time and the day of the week, it is possible to infer the 2 most significant places – user’s home and work. The information that a person every morning commutes from one place to another, spends around 8 hours in the second places and then go back to the first
one, implies that those two locations can be identify as a home and work. This fact is easier to infer when the observed individual has predetermined constant working hours. In this case it may be also fast concluded when the person is on holidays (user sleeps in a new place and is not going to work during the week). Another sample of using the knowledge that understanding patterns brings may be recognizing people around (detecting other mobile phones). After analyzing places and time of the day it is possible to detect who is a part of a family or a flat mate (person noticed mainly at home and almost always visible during night hours). Additionally, co-workers can be identified as people that are near only at predefined work place, friends can be noticed at evenings and during weekends, etc.

The final part of the 3 level model of context is wisdom. That is, understanding of the principles, inferring from the gather knowledge and building the intelligence. The examples can be deducing that a meeting will start later if one of the buddies is delayed or realized that a person moved to a new house (different district/street) when he spends night in different place but still commute to work in the morning.
Because of a variety of daily schedule and people’s behaviors, some information can be incomplete and the collected data cannot be generalized. Also all the noise, that is captured by sensors, has to be taken into consideration. The automatic analysis of real-life from mobile sensor data is still a challenging area.

2.2.2 The role of sensors

By carrying a modern mobile phone a human being becomes the focal point of sensing (11). Creating a personal sensor network, where the device owner turns into a sensor hub (12) and visualization of the collected information can bring benefits not only for the monitored person but also for his friends.

Examples of usage of the selected sensors along with existing services sample are presented below.

Localization

Numerous technologies, offered by mobile phones, can provide this contextual information (GPS, GSM (cell tower identifiers, WiFi access points).

Data collected by the GPS are the most precise ones but also require a specific amount of time and the GPS module on the mobile phone must be turned on. Coordinates (latitude, longitude, and altitude), defines the location. They have to be converted into meaningful information (e.g. point on the map).

In areas such as office buildings, campuses or individuals’ homes the information about a physical location may be available thanks to a WiFi access point. Additionally a second type of data can be taken into consideration, such as: access point identifier and signal strength.

The cell phone’s location can be also verified by the Base Transceiver Stations (BTS), which are used for relaying the phones’ data to the network. Based on strength of the signal it is possible to determine which local transmission tower is the nearest. The information from a GSM cell ID is not as accurate as the one from a GPS module because it relies on distribution of the stations (around 500 meters in urban regions, up to 15 km in rural zones) (13). This data is still sufficient to estimate the user location (accuracy rises when the phone is in a region with three or more base stations in range (14), presented on Figure 2).

---

2 Global System for Mobile Communications
Finding the user’s whereabouts is a technical challenge. The bigger problem is to utilize this information as much as possible. After the “points on the map” where a user has been located, an application developer can make headway in solving the problem of identifying significant places for the user. In this case a human can be considered as a “set” of several identities (15), where each identity (role) is composed of characteristic meanings and behaviors. Usually the place, where a user spends lots of time or stay the more often, is related with his current role (worker, father, movie lover, etc.). Similar analysis can be done for behaviors’ patterns (e.g. work often follows home).

Another model may be created by an observation of changes in a user’s daily life. For example, during summer months (e.g. June), a mobile phone is far away from its usual location. The most obvious explanation seems to be that the user is on holidays (especially if the current place is a foreign country).

In addition, individual who likes sushi can receive a text message saying that a sushi restaurant is nearby. Message can be very useful especially during the lunch time.

**Movement**

Accelerometer is a sensor that may be used to detect user’s motion and at the same time diagnose the activity (identify if the phone is moving or lying idle). By knowing the speed of the moving devise it is possible to deduce if user is walking, going by car/bus, riding a bike etc.
Nearby devices

Discovery protocols such as Bluetooth and WiFi are very often used to scan for other co-located devices. That entails information about surrounding people. All Bluetooth enabled devices have a MAC address (that is unique) and a name (determined by the user). A set of discovered MAC addresses and names, after eliminating devices that are not phones\(^3\), can be mapped into a group of people that are close to the user. Presence-sharing services can base the identity analysis on the mobile phone’s Bluetooth name. A simple scheme may assume that a user includes an email address or his own name in the device name.

Very often the location is closely related with the surrounding people (e.g. family is at home, at co-workers are at work). The nearby devices detection can also allow users to find people sharing similar interests. An application on the phone can scan the surroundings, analyzing the identities of co-located users and point a person who also is looking for new friends. Another example may be using this technology during technology-oriented conferences. The mobile phones can inform users about proximity of participants who work on a related field of study (by sending an alert with a photo and name of the person).

Audible Noise

One of the less obvious sensors might be a microphone built in the phone. People behind the project NoiseTube\(^4\) decided to utilize the potential of sound recording. NoiseTube is a research project which tries to approach the problem of “noise pollution”. Its aim is to convert each of the mobile phones into a noise sensor (12). Those are to be utilized by people living in crowded and noisy cities, thus enabling them to measure the noise levels in dB(A) they are exposed to. End to end project’s ultimate goal is to create a collaborative map of places, where people are more exposed to harmful influence of the noise (e.g. hearing loss, heightened stress, sleep interruption, decrease school performance). The map is created on the basis of geolocalized information (GPS) and sound records picked up by the mobile phone. All the collected information is send to the NoiseTube server and convert into a Google Earth\(^5\) compatible format. SoundSense is another project which focuses on the sound recording (16). The application implemented on the Apple iPhone is capable of recognizing the most significant sound form in human’s life. The information recorded by the phone’s microphone is processed and categorized. SoundSense tracks the user behavior by monitoring the everyday sounds and

\(^3\) Determined on the basis of a „Class of device“ parameter (36)
\(^4\) http://noisetube.net
\(^5\) http://earth.google.com
classifying their types e.g. talking, listen to the music, walking outside, brushing teeth, cycling, driving in the car, etc. (17).

More information about SoundSense and other sensor based application can be found on MertoSense Project\(^6\) website.

**Data from the phone**

Current profile and phone’s battery status also should be mentioned. Among all the data captured by sensors and required more or less advanced analyzes, these two pieces of information seem to be the most obvious and easy to capture. Knowing that somebody’s phone has only 1% of the battery (and it’s not charging), the first thought is “his/her phone is going to die in a moment” therefore there is no motivation for calling. The same situation is related to the current profile, especially the “Silent” one. It happens that the person we are calling does not pick up the phone. The profile knowledge makes it obvious, that you cannot answer the phone when you don’t know it is ringing (because you don’t hear it).

Additionally, a calendar accessible in the phone can bring more details about user’s daily schedule. Information about planned meetings, birthdays’ dates, important things that should be remembered, all this things may be obtained from the calendar.

**Disadvantages and limitations**

There are disadvantages and limitations of the mentioned sensors. The GPS weaknesses can be buildings, trees and glass coatings in cars and trains. All these factors can cause signal shadowing and make the GPS module almost useless. Another disadvantage is the huge battery consumption that accompanies the GPS measurements.

The information from GSM cell identifiers assures a good indoor reception, but it does not allow separating significant places which are near each other, i.e. that are covered by the same cells. WiFi access points can be used for positioning only, when the exact location of access point is known and when it’s possible to associate the network name with the place.

Bluetooth can also be problematic when it detects other Bluetooth enabled devices that are not in visual proximity (e.g. people with phones that are behind the walls).

\(^6\) [http://metrosense.cs.dartmouth.edu](http://metrosense.cs.dartmouth.edu)
Accelerometer, it can help to recognize user’s activity, but it never will be as accurate as the manual user’s input. For example, a phone can recognize if a person is going by car but it cannot conclude as a driver or passenger.

Finally, the most obvious one, a phone can be left behind, and recording data that are not related with the user.

2.3 The Social Networking Sites phenomenon

The revolution of the Web 2.0 brought a new look at people’s social life. The Internet users were invited to contribute and interact on the Web sites. This novelty resulted in creating different kinds of Social Networking Sites, which offered a variety of possibilities to make new friends or business contacts, to share knowledge or simple life experiences and adventures. SNS have been popular from the beginning of their existence. They bring opportunities for people to start sharing content and personal information about themselves, and sharing is a starting point for communication (8).

2.3.1 The most popular Social Network Site

All SNSs usually offer the same basic functionalities: the friends’ listings, person searching, private messaging, discussion forums, events management, blogging, commenting, and data uploading. The differences between SNS can be found in the meaning of users’ relationships. It is a reason represents the reason, why this kind of network has been chosen. Let’s take a closer look at some of the most popular virtual communities.

Facebook\(^7\) is a sample of SNS that attracts more and more people. With the passing years it has reached over 400 million active users and the numbers continuously grow (18). That makes it the world’s biggest Social Network. With Facebook, which was originally designed only for students of the Harvard College, the meaning of “friends” seems to be natural. Users, who declare themselves to be friends, usually knew each other before in the real life.

Facebook biggest rival, MySpace\(^8\), represents a weaker meaning of “friendship” connections. On this site, contacts usually have never met. MySpace seems to be dominated by teenagers who want to boost their social ego. While finding a celebrity, an actor or a singer on an ordinary men’s friend list, it’s easy to deduce that MySpace is also a great place for artists to get closer with their fans (8).

\(^7\) http://www.facebook.com/
\(^8\) http://www.myspace.com/
The next type of a social network is LinkedIn\(^9\), which can be perceived as a professional one. LinkedIn helps to exchange information and maintain business contacts with others. Moreover some users simply gather as many contacts as possible, just to increase their visibility to potential employers (8).

Twitter\(^{10}\) is a different type of SNS, where contacts are called “followers” and are notified of other user’s social updates. The posted messages are up to 140 characters long and are displayed on the author’s profile page, where users who monitor the profile can see them (8).

**Comparison**

When analyzing the mentioned services from a developer point of view, the most important criteria is an API\(^{11}\) that provides an easy and quick access to the network. The Table 1 presents a summary of the provided APIs (19).

**Table 1: Summary of provided APIs**

<table>
<thead>
<tr>
<th>API</th>
<th>Features</th>
</tr>
</thead>
</table>
| Facebook | public API  
|        | RESTful\(^{12}\) protocol  
|        | allows to use the social connections and profile information: users can add social context to their applications by utilizing profile, friend, Page, group, photo, and event data, privacy settings available  
|        | supports XML data format  
| MySpace | public API  
|        | MySpace Developer Platform (MDP)  
|        | created applications integrate directly into MySpace pages and get exposed to everyone in the network  
| LinkedIn | available if developer passes an approval process  
| Twitter | public API  
|        | two RESTful APIs: REST API methods allow to access core Twitter data (update timelines, status data, and user information) Search API methods give possibility to interact with Twitter Search and trends data  
|        | supports XML and JSON data formats |

---

\(^9\) [http://www.linkedin.com/](http://www.linkedin.com/)

\(^{10}\) [http://twitter.com/](http://twitter.com/)

\(^{11}\) Application Programming Interface

As LinkedIn has been created for business oriented users and Twitter is more like a microblogging service, those two networks will be excluded from the further analysis. An in depth comparison (Table 2) will be performed for the Facebook and MySpace systems that are the two largest social networks.

<table>
<thead>
<tr>
<th>Table 2: Facebook and MySpace comparison (20)(21)(22)(23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
</tr>
<tr>
<td>Launched</td>
</tr>
<tr>
<td>How users presents themselves</td>
</tr>
<tr>
<td>Creating profile - background customization</td>
</tr>
<tr>
<td>Relationships</td>
</tr>
<tr>
<td>Groups</td>
</tr>
<tr>
<td>Keeping track of news</td>
</tr>
<tr>
<td>Messaging</td>
</tr>
<tr>
<td>Music uploads</td>
</tr>
<tr>
<td>Photos uploads</td>
</tr>
<tr>
<td>Video uploads</td>
</tr>
<tr>
<td>Share media</td>
</tr>
</tbody>
</table>

<sup>13</sup> http://photobucket.com/
<sup>14</sup> http://www.adobe.com/flashplatform/
<table>
<thead>
<tr>
<th>links</th>
<th>Content restrictions</th>
<th>Independent applications</th>
<th>Security</th>
<th>General idea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>user restricted from uploading “unwelcome”(^{15}) content</td>
<td>fewer content restrictions(^{16})</td>
<td>more independent applications created by 3(^{rd}) party developers</td>
<td>default privacy settings account are limited to a users list of “friends”, easy to hide info from certain people, possibility of keeping some information private</td>
</tr>
</tbody>
</table>

The differences between Facebook and MySpace may be hard to notice. Both services are social networking sites that have become addictions in our society. However, Facebook seems to be more secure and the idea behind it gives the impression that members represent a real community.

### 2.3.2 Bringing context-awareness into the Social Networking Sites

With passing time, many SNS become meaningless, as their providers are neither successful in keeping users’ attention nor able to let them improve the network. A reasonable solution of this problem would be to bring a context-awareness into the SNS. A key role in this process is played by mobile devices, particularly mobile phones. The already existing convergence of mobile technology and the “traditional” Internet application enables the creation of a personalized profile of the social network user and at the same time improves the SNS and makes it more interesting and valuable. A voluntary user input can be partially replaced by one that is automatically generated by the sensors. Sensors enable various environmental monitoring and at the same time provide contextual data. Social Network Sites can be used just as a “storage infrastructure” for this data.

Examples of usage of the selected sensors along with existing services sample are presented below.


Location

It seems that in context-aware applications the user’s location data is the most widely analyzed and processed information (15). Presenting the current location on SNS is a sign for the user’s friends saying: “look I’m here”, but how the user can benefit from this place disclosure. The “profit” can be visible for example during traveling. SNS members can share their recommendations. For example, when someone is touring a new city, by posting this information he makes easier for himself to obtain an interesting data (e.g. museums to visit, good restaurants and other local tips) from friends who have already been there. Another possibility, individual who likes sushi can receive a text message saying that a sushi restaurant is nearby. Message can be very useful especially during the lunch time.

There are numbers of services existing on the Internet which allow tracing the location of a cell phone. One of them is CenceMe\textsuperscript{17}, a personal sensing system that offers the possibility to present current locations. The user is able to automatically export his whereabouts information to the SNS (e.g. Facebook, MySpace) and control a historical analysis of his activity (e.g., how often did I go to the gym this week?) (24).

Another mobile application, developed in Python and running on Nokia Series 60 phones, is IYOUIT\textsuperscript{18}. The goal of the application is to automatically record, store and use context information (share with others, do statistics, etc.). Apart from locations (GPS and GSM cell based) Context Watcher records also moods (based on user input), activities and meetings (based on reasoning), body data (based on heart and foot sensors), weather details (based on a location-inferred remote weather CP), visual data (pictures enhanced with contextual data) (15).

The user’s location can be presented in various ways. The type of information that the SNS member wishes to present depends on his motives Some people may want to disclose the exact location (address) or define the place in general (home, work or airport) other may just want to suggest to their friends the approximate area (far away from home, different city, different country).

Presence – sharing

Presence-sharing, possible thanks to the Bluetooth interface, can inform a Social Network members if their friends are around. The information, who user met, may be also send to the network and displayed on user’s profile.

\textsuperscript{17} http://www.cenceme.org
\textsuperscript{18} https://www.iyouit.eu
Data related with proximity, after analyzing it could help in determine the relations between social network users (showing who is a part of a family, who is just a co-worker).

**Activity**

By sharing the activity on social network site, members have possibility to inform friends if they are busy or not. Often the activity is related to the current location of the user and its presentation can be a way to encourage friends to join.

**Sounds**

The data related with the detected noise can also be utilized in SNS (e.g. showing if the user is listening to music or sitting in a quiet place). All the “noise logs” (gathered via GPS or via Wi-Fi positioning techniques), recorded by the network member and posted on his profile, can also help others to find a “calm place” preferable for example for a business meeting.

**Data from the phone**

Social Network Site can be a place where its members can go and check if their friends are available at the moment. The information provided by the phone may show the future plans of the user or just explain why the person we are calling does not pick up the phone.

### 2.4 Privacy issues

Many SNS are criticized for privacy issues because most of their users tend to expose frequently personal (and even intimate) content on these communities. The content includes material status, date of birth, phone number, job, location and email address. Has nowadays a lack of privacy online become the norm?

Mark Zuckerberg, the founder of social networking giant Facebook, believes that people no longer expect their details to be private. He claims that putting more and more information about our life online is a new trend (25). He reminisces the year 2004, when Facebook had its beginning: “When I got started in my dorm room at Harvard, the question a lot of people asked was: ‘why would I want to put any information on the internet at all? Why would I want to have a website?’” (25) Nowadays the Social Networks’ users openly share different kind of information with significant amount of people.
2.4.1 Private data presented online

An intriguing question come to mind - what is so extraordinary about social services, that users are willing to give away their private information inside a network. Why are we showing our life to our 100 “Facebook friends”, whom sometimes we have never met, while we wouldn’t do it with a random person met on the street? What is it about Facebook that makes us put so much personal information?

Most humans want and need to interact socially, and that’s why they voluntary ‘scarify’ their privacy. Very often the social pressure is so strong, that it seems to be no other way to keep in touch with others. “If you’re not on MySpace, you don’t exist” or “I created an account, because otherwise my friends would have done it for me” (26). Some people (especially teenagers) feel that they have no choice but to join or be left out socially.

The value of the SNS is growing along with the increasing numbers of users and the content that they are bringing. This is the reason why SNS providers are encouraging people to use their platform more intensively. They are offering to the network’s members numerous possibilities to present information (e.g. blog, wall, pictures, uploading date through mobile phone). People utilize all the new tools and increase the amount of data on the SNS profile. They are doing it with pleasure because the richer the profile, the more interesting for other SNS members. (27)

Networks such as Facebook can be considered as tools that makes everyone a bit more narcissistic. SNS allow users to present themselves to the world in the way that they want to be seen. Some of them are uploading only the best photos and write only positive comments. Others just want to boast about a new thing they have just purchased. This process is becoming increasingly common and aggressive. Telling to the Internet that e.g. “John Smith spent $3000 at the BMW dealership” is not a thoughtless idea. It is already happen in the US. The social media “sharing” site called Blippy\(^{19}\) allows the users to post their favorite purchases from e.g. iTunes, Amazon, Visa, MasterCard and more (28).

2.4.2 Threats

We use SNS to link up with friends, but for some reason we forget that information is not private. The virtual communities help people to interact. Going through one’s “profile” we can learn more about his/her interests and activities. Unfortunately, the fact is that not everybody is searching for a friend. When we are talking with someone in the real life, we are trying to moderate the conversation depending on who is our interlocutor. Online the boundaries are

\(^{19}\)http://blippy.com/
gone. As it’s very easy to become a ‘star of the Internet’ for the wrong reasons, SNS’ users have to think twice before posting anything.

It’s possible to distinguish three types of farms which can take place when the privacy protection is neglected (27):

- **Information-base harm** – easy access, on the profile, to email address or mobile phone can be a cause of harassment by others
- **Information inequality** – data about what the users' buy (thus where they spend their money) and their preferences / habits could be utilized in the marketing area and price shaping. All of that without alerting users' awareness
- **Information injustice** – uploaded information, for example party photos, that seems to be funny for the friends, bring the risk of using it to compromise the user (e.g. while applying for a job)

When consider the types of threats we can try to specify who wants to harm the SNS users. In this purpose we need to find out who has access to the data that are stored by the SNS. The types of attackers are listed below (27):

- **Other users** – people who also have an account on the SNS. Usually users can choose if they want their profile to be “public” or “private”. If the account has been marked as “public” everyone has access to the information left online. Otherwise, the user can decide who of the SNS members is allowed to see the full profile page and depending on the connection between the contacts, how much of the private information is presented.
- **Third parties** – people or organizations that have no user account on the SNS. Instead of having full access to the SNS, they can only view the limited information (the publicly available data).
- **Platform providers** – operators and owners of the SNS. The operators have full access to the complete information that users uploaded. Another thing is that, the data can be stored on the SNS servers for an indefinite period of time.

There are various reasons why people violate others privacy. We monitor our friends’ profiles on SNS mainly because of curiosity. The attackers however have other motivations – social and financial. Mainly the application provides and the third parties receive profits from disclosing the
private data, collected on the SNS, to the marketing companies or personal credit rating agencies. That cases the SNS users to be harassed by different types of advertising (27).

The location-based services on SNS can raise users’ biggest doubts. Possibility of posting the current location can cause a lot of damages. Parents would be able to spy on their children, bosses have opportunity to check, what the employees are doing in their free time or on the business trips and the government has an easy way to keep the inhabitants under surveillance. Even information about hobbies or visited places (pubs, restaurants, events), that seems to be irrelevant, can be a reason of harm (13). Other people’s judgment very often hurts the potential victim’s fillings.

Numerous of people are unaware of the fact, that theirs mobile phone are constantly generating location data and the only way to avoid this privacy violation is to switch the device off. Right now the LBS present the collected information only if the user gives his content. Unfortunately the LBS providers still have access to the location data and can make, even unintentionally, a match between the profile and the places where a user has been (13). There are many situations people would like to keep, avoiding others to speculate over.

The short-range wireless technology, such as Bluetooth, allows users to find all the Bluetooth-enabled devices. This technology became a standard on most of the mobile phones and the possibility of identifying people around us can be easily abused. The presence-sharing should be accessible only if all the present people approve it.

The same problem concern the data captured by the phone’s microphone. People are sensitive about how audio records (especially a conversation) are used by the system.

2.5 Summary

Mobile phones became an essential part of human’s life. This fact helped to conduct an experiment, which proved that people are not as unpredictable as they thought. The chapter 2.1 explained how it is possible, that the whereabouts of average person can be predicted up to 93% of the time. The modern mobile phones, equipped with sensors (GPS, accelerometer, microphone, etc.), described in chapter 2.2, allow turning human into a focal point of sensing. The gathered information about whereabouts, people around, profile of the phone, battery state and many more, represents user’s context. This data can be shared on Social Network Sites to enrich the members’ profiles and show availability to the friends.

By comparing the presented in chapter 2.3 Social Network Sites, we can notice that Facebook is mainly used for “social searching” (i.e. looking up already known people), while MySpace is more
like “social browsing” (i.e. finding strangers online) (8). LinkedIn has been created for business reasons, not for entertainment. All the SNS mentioned above require a mutual approval of the connection as “contact”/“friend”, Twitter however does not. The lack of “friend” approval results in none privacy protection, which disqualifies Twitter as a place, where most of the people want to present confidential information.

After this brief analysis, Facebook seems to be the best choice for the realization of the practical part of the thesis. Control over who can be you friend and what that friend is able to see of your profile, provides grounds for trusting the SNS. The approbation from both parties and additionally acquaintance in the real life, enforce a sense of privacy protection. If we want to display our private information online, it is important to control whom it is shared with.

To improve social communications a contextual knowledge can be utilized. The information obtained by phones’ sensors can be translated into various social scenarios, for example contact proximity notification, activity proposal, trip planning assistance. Unfortunately the most precise information (especially about activity) can be only given by human. Sensors available in the mobile phones make it easier to collect data about users’ context. Users can analyzed the collected information, log it for future statistics or just share it with friends on the SNS and at the same time share a part of their private life. Unfortunately a cell phones that listen, learn and present exact and correct data is still a big challenge.

The chapter 2.4 brought up the privacy issues topic. The amount of the online published information is growing fast. Social networks’ providers are pushing users to expose increasingly personal details and they are keeping all the data as long as possible, because the value of the network is in its size. SNS will continue to exist adapting themselves to the consumer needs, which are changing with time and developing technology.

The today’s pressure of following the trends and being “in touch” with friends all the time causes, the „social” thinking seems to take the precedence over the “logical” one. The best example is Blippy.com, the “Twitter of personal finance”. All the types of social services leave enormous commercial opportunities and those opportunities will be exploited, with or without users’ approval. Even though SNS’ members are aware of the potential harm, they still reveal confidential and embarrassing information about themselves. They prefer the social benefits over the privacy infringements. However, if ignorance is the reason why SNS users do not care about their privacy, they should be educated about the risks.
3 Design

In the previous chapter context definition along with the mobile phones’ sensors utilization have been presented. Basing on this information it is possible to design an application that uses mobile phone sensors and presents the gathered data on the Social Network Site.

This chapter will describe what aspects should be taken into consideration while designing an application for a particular SNS platform. Then the general architecture of the application on the phone as well as on the Social Network Site will be explained.

3.1 General idea

There are plentiful of Internet services (29) that are based on the data received from the mobile phones. For example, Whrrl20 and Rummble21 are location-based social networks that work both on the mobile phone and the classical web frontend. Mig3322 mixes a Web-based social network with a cell phone technology. Next are Loopt23 and Utterli24, services pushing mobile-generated content to the Web and keep people posted all the time. Another one is GyPSii25, a location-based social-networking platform designed for mobile devices.

The biggest disadvantage of this kind of services is that they are quite new and not well known. The majority of SNS members do not want change their habits. They already have a profile on “the old” network and they put a lot of effort into it. They spent a lot of time collecting friends and getting to know the social network that they have already chosen. Users do not want to start everything from the beginning (create a new profile, inform everyone about the new SNS, etc.). More attractive for the members, is to change the well know social network and enrich it by utilizing the possibilities that mobile phones offer.

3.2 Assumptions

Mobile users are almost inseparable from their phones. The device goes with people all the time and it doesn’t matter if it is a smart phone with all the fancy features or just a traditional phone. As those devices evolve they are no longer solely utilized as phones. They begin to extend their users’ personal computers, entertainment systems and daily information sources.

---

20 http://whrrl.com/
21 http://www.rummble.com/
22 http://mig33.com/
23 http://www.loopt.com/
24 http://www.utterli.com/
25 http://www.gypsii.com/
Access to the Internet through the cell phone became something natural. Millions mobile subscribers surf the web (30) (31). They check emails, make bank transactions, look at social network sites and read news on their favorite websites. The statistics show that number of mobile Internet subscribers is steadily rising. That is due to reduced cost of service, complexity and improved users’ experience (30).

Assuming that users are carrying their phone all the time and they have access to the mobile Internet, it is reasonable to build mobile services that utilize phones possibilities. The services can both entertain the users and make their life easier.

### 3.3 ContextBox application

The goal of the application is to show the basic information about a user’s status without forcing him to type it manually on the phone and send it to a Social Networking Side. The application should be transparent from the user’s point of view.

The “ContextBox” system architecture is presented on Figure 3.

---

**Figure 3: ContextBox system architecture**
The system is composed of 3 parts:

1. Client application working on the phone that is responsible for collecting information from the sensors.
2. Google App Engine server responsible for hosting the whole application and analyzing the data.
3. Facebook platform responsible for presenting the data to the users.

The first tier represents the mobile phone equipped with sensors. Periodically the phone has to send a report to the backend (server). Before submitting anything the phone client application determines if there were any changes in the context information since the previous submission (do not send the same data twice e.g. the same profile). The data is serialized and sent to the server. The tier 2 represents the Model as well as the Controller in the MVC\(^{26}\) architectural pattern. The database is responsible for storing the gathered information. After the system analyzes the phone’s report and saves the newest data, the application on the Social Network profile is updated. The tier 3, which is a user’s profile, represents the View in the MVC pattern.

### 3.3.1 Collecting data from phone’s sensors

Basing on the multilevel context model, presented on the Figure 1 in chapter 2, the raw sensor data represents the bottom layer – information. Description of the functional requirements of the phone application, that allows meeting the quantitative demand for information, is presented by means of an UML Use Case diagram. Figure 4 shows the relationship between use cases within a system and their actors. Additionally Table 3 contains a detailed description of a base Get sensor data use case.

#### Table 3: Get sensor data use case

<table>
<thead>
<tr>
<th>Use case name</th>
<th>Get sensor data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>A phone’s application – mobile Context Toolbox (MCT) is responsible for collecting sensor information.</td>
</tr>
<tr>
<td><strong>Preconditions</strong></td>
<td>A user has installed the application on his/her phone. The phone has access to the Internet and is equipped with Bluetooth, accelerometer and GPS modules.</td>
</tr>
<tr>
<td><strong>Actors</strong></td>
<td>Phone MCT</td>
</tr>
</tbody>
</table>
| **Steps** | 1. Application requests checking the status of the sensors  
2. REPEAT:  
   - Get profile phone  
   - Check the battery state  
   - Get sounds samples (microphone records)  
   - Detect movement (accelerometer indicator) |

\(^{26}\) Model–View–Controller
- Check current calendar events
- Detect nearby Wi-Fi (get name and MAC address)
- Detect nearby Bluetooth enabled devices (get name and MAC address)
- Get GPS coordinates (latitude, longitude)
- Get Cell ID information

**UNTIL:**
- All sensors checked

**Post conditions**
The gathered data is stored and sent to the Google App Engine server for processing.

**Notes**
Fetching sensors’ readings occur in regular intervals.

The collected data should be stored on the phone as long as there is possibility of sending it to the server for future analysis.

![Figure 4: Information level Use Case Diagram](image)

The core of the phone application will be the Context Toolbox (32) platform. The Content Toolbox is an environment developed as a master project on the Technical University of Denmark (DTU). It enables rapid application prototyping for the mobile phones. The platform was built with a view to speed up the process of context aware systems development. Context Toolbox will be responsible for collecting the data from phone’s sensors.
The multilayer architecture of the Context Toolbox is presented on Figure 5.

The application starts automatically when the phone is turned on. It should work in the background, so the interaction between the user and the application must be reduced to the minimum. Some of the Context Toolbox’s options had to be excluded with the aim of decreasing the battery consumption (e.g. GPS and accelerometer readings).

The Context Toolbox application has to be adapted to the server layer, so some parts of the code have to be converted.
The data that will be obtained from the user’s phone is listed below:

- Phone’s battery state (in %)
- Phone’s profile (along with sound on/off information)
- General information about the surroundings collected using the Bluetooth interface (people around/user is alone)
- Current location obtained from the GSM / UMTS Cell IDs

The battery state will be shown in a percentage scale to signalize the possibility of the phone shutting down (very low battery state).

The profile of the phone explains the current situation of the user. By providing the information about the silent state, the system will inform potential callers, that it’s not the best time for having a conversation.

The data collected using the Bluetooth interface will give the general picture of the user’s surroundings. The system will eliminate the insignificant devices such as printers, computers, etc. and focus only on the number of detected phones (determined on the basis of a “Class of device” parameter). This information (combined with the location data) will signalize if user is sitting alone (and maybe needs a company) or is surrounded by people.

Current user’s location will be obtained on the basis of the GSM / UMTS Cell ID information. This allows to imply if the phone’s user if far away from home. Technique does not need any interaction with the user.

All the data gives the general image of the user’s state (available or busy). The system sets the battery economy and the user’s independence above the precision of the location. The GPS coordinates are the most accurate. Unfortunately this technique of finding the user’s whereabouts is not the best choice for a long term data gathering with the phone (the battery consumption rate is unacceptable). Also tagging (@home, @work, etc.) detected places by the user would be the fastest way for determining his location. However this could be annoying and meaningless as the assumption for the application is to not to force any manual input on the user.

3.3.2 Hosting the application and storing the data

Thanks to an external server, that hosts the application, there are almost no limitations in the amount of data that can be captured. Practically there are no limitations of the processing
power, which can be noticed on the phone, because all the calculations take place on the server. For hosting purpose Google’s infrastructure has been chosen. More about the server platform will be presented in Chapter 4.

The server side of the system is responsible for processing data that is obtained by the mobile phone. The top two levels of the context model (Figure 1), wisdom and knowledge, are embedded in this part of the system. The application on the server will be responsible for storing and processing the sensor data. The UML diagram, presented on Figure 6, shows the relationships between individual use cases. The base use case (green color on the diagram) represents the final results of the data analysis – the principles understanding, that can be achieved by means of combining all of the related knowledge.

Figure 6: Knowledge (blue color) and wisdom (green color) levels Use Case Diagram
Tables presented below briefly describe the utilization of information collected by the sensors.

**Table 4: Infer meaningful places use case**

<table>
<thead>
<tr>
<th>Use case name</th>
<th>Infer meaningful places</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>System can diagnose what locations are meaningful for the user (home, work, school, etc.)</td>
</tr>
<tr>
<td>Preconditions</td>
<td>GPS coordinates or Cell ID information has been sent to the server by the mobile phone.</td>
</tr>
<tr>
<td>Actors</td>
<td>Google App Engine</td>
</tr>
</tbody>
</table>
| Steps           | 1. Take current location  
                  2. New location is placed on the map  
                  3. REPEAT:  
                  - Monitor the duration of a stay in a place  
                  - Monitor the time of the day  
                  - Monitor the frequency of visiting the place from 2.  
                  UNTIL:  
                  - New location detected  
                  4. Analyze the collected information, compare with previous readings |
| Alternative path| 2. Location already known  
                  3. Place location on the map  
                  4. Tag location with a proper name if possible (home, work, etc.) |
| Post conditions | The current user’s location can be displayed on the map. |
| Notes           | In order to infer the meaningful location a long term analysis has to be performed. |

**Table 5: Infer if user is busy use case**

<table>
<thead>
<tr>
<th>Use case name</th>
<th>Infer if user is busy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>System can diagnose the availability of the user</td>
</tr>
</tbody>
</table>
| Preconditions   | The phone’s profile information has been sent to the server.  
                  Optionally, the meaningful location of the user is known. |
| Actors          | Google App Engine |
| Steps           | 1. Take the phone’s profile setting (silent, general, etc.)  
                  2. Detect if speaker is muted (on/off)  
                  3. If possible connect profile with location |
| Post conditions | The current user status (can/cannot pick up the phone) can be inferred.  
                  Further on automatic profile adjustment to the current context can be made. |
| Notes           | The user’s availability can be deduced on the basis of his location (e.g at work – busy profile). |
Table 6: Infer activity use case

<table>
<thead>
<tr>
<th>Use case name</th>
<th>Infer activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>System can diagnose the user’s activity</td>
</tr>
<tr>
<td>Preconditions</td>
<td>Detected motion data has been sent to the server. Optionally, the meaningful location of the user is known.</td>
</tr>
<tr>
<td>Actors</td>
<td>Google App Engine</td>
</tr>
</tbody>
</table>
| Steps          | 1. Check for movement information  
                 2. User does not move  
                 3. If the meaningful location or relations with detected people are known try to infer the activity |
| Alternative paths | 2a. Moderate movement  
                        2b. Significant movement  
                        3. If the previous meaningful location is known try to infer the activity |
| Post conditions| None |
| Notes          | Sample: movement rate — still, meaningful location — London pub, detected people — John, Mary the result of analysis — partying, socializing |

Table 7: Infer relationship type use case

<table>
<thead>
<tr>
<th>Use case name</th>
<th>Infer relationship type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>System can diagnose the relationship between the user and an encountered person.</td>
</tr>
<tr>
<td>Preconditions</td>
<td>Detected Bluetooth enabled phone’s MAC addresses have been sent to the server. Optionally, the meaningful location of the user is known.</td>
</tr>
<tr>
<td>Actors</td>
<td>Google App Engine</td>
</tr>
</tbody>
</table>
| Steps          | 1. Search a data base  
                 2. New MAC address saved in the data base  
                 3. Increase a meeting counter  
                 4. REPEAT:  
                    - Monitor information about the location of the encounter and time of the day  
                    - Monitor information about the duration of the encounter  
                    UNTIL:  
                    - End of the encounter |
| Alternative path | 2. MAC address found in the data base  
                        3. Take a person’s name  
                        4. Increase a meeting counter  
                        5. REPEAT:  
                           - Monitor information about the location of the encounter and time of the day  
                           - Monitor information about the duration of the encounter  
                           UNTIL:  
                           - End of the encounter |
| Post conditions| The person’s name will be known during the next encounter, if the new MAC address is saved in the data base. The location of the meeting allows defining the type of the relationship. |
| Notes          | In order to infer the relationship a long term analysis has to be performed. |
Table 8: Present information use case

<table>
<thead>
<tr>
<th>Use case name</th>
<th>Present information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Context data can be displayed on social network site’s member.</td>
</tr>
<tr>
<td>Preconditions</td>
<td>The information is processed and connected to the proper social network site’s member (all data tagged with the user’s ID).</td>
</tr>
<tr>
<td>Actors</td>
<td>Google App Engine, SNS</td>
</tr>
<tr>
<td>Steps</td>
<td>1. Updated context data displayed on the profile</td>
</tr>
<tr>
<td>Post conditions</td>
<td>None</td>
</tr>
<tr>
<td>Notes</td>
<td>User must be able to access privacy settings of the application.</td>
</tr>
</tbody>
</table>

3.3.3 Social Networking platform as a context presentation layer

Allowing developers to create programs, which can appear as public widgets on user’s profile, was a huge breakthrough in SNS world. This possibility results in growing interest in this kind of networks and allows them to “live its own life”.

![Figure 7: Typical user profile site. The profile contains: place for a SNS member’s photo [1], general information about a SNS member [2], a contacts’ list [3], that SNS member can personalize [4], SNS member’s name and correct status [5], settings list[6], “Wall”, place where a SNS member and his/her contacts can write comments [7]](image-url)
The SNS will be used as a presentation layer for the application. Figure 7 depicts a general design of user profile in a Social Networking Site. The constant areas of each profile are:

- 1 - place for a SNS member’s photo
- 2 - general information about a SNS member
- 3 - a contacts’ list
- 5 - SNS member’s name and correct status
- 6 - settings list
- 7 - “wall”, a place where a SNS member and his/her contacts can write comments

The area 4 represents a space which a SNS member can personalize (e.g. display additional applications).

The region 4 will be the best place for presenting the user’s context. It is built in the profile and it is conspicuous. Unfortunately the available space is limited. This is a reason why during the implementation it is necessary to focus on presenting the context information in a way that does not confuse the user.

A small box on the profile will be purposed just to depict the general state of a SNS member (stating if the user is available e.g. alone and the General profile is set on the phone or it’s not the best time for calling e.g. Silent mode of the phone). More detailed context information can be displayed on a separate page, which will be available after installing the ContextBox application on the SNS.

### 3.4 Communication flow

Figure 8 presents an information flow sequence diagram.

The application will start automatically, simultaneously with turning on the phone. The user does not the take any actions in order to send the information to the server. The application on the phone is systematically searching for any changes in the user’s context. If any modifications (comparing to the previous state) are detected, the phone is sending a new report with the most current data. Then after the data is processed by the server, the user’s profile on SNS is updated.

The data submission process between the phone and the server is unidirectional. The server is only sending an acknowledgement confirming that the data has been received. No other information is sent from the server to the client phone application.
The prototype of the system is ignoring all of the errors that can occur during the transmission (timeouts, server’s inaccessibility, etc.). If there are problems with sending the report, the phone is waiting for the next opportunity to do so.

In a production system a communication exception handling should be implemented.
4 Implementation

The requirements and design of the system have been presented in the previous chapter. In this section implementation of the system will be described.

As it was indicated in Chapter 2, Facebook has been chosen as a Social Network platform for the context aware application.

4.1 ContextBox components

The main two components of the system are:

- Mobile Context Toolbox (MCT) mobile phone application, responsible for collecting information from the sensors and sending the data to the server.
- Google App Engine service, responsible for storing, processing and relaying the data to Facebook.

The mobile application is a modified version of MCT, developed in Python and running on the Symbian v9.2. For the hardware platform Nokia N95 8GB smart-phone has been chosen (S60 3rd Edition, Feature Pack 1)27. As mentioned above the development environment is Python for S60 version 1.4.528. Detailed description of the MCT phone application is described in the separate thesis (32). In this paper only the implemented modifications will be explained.

Google App Engine (33)(34) allows developers to run web applications on Google’s infrastructure. The service supports Python runtime environment (up to version 2.5.229), which includes Python interpreter and the Python standard library. App Engine also provides a Python web application framework called Django30 which takes care of the basic request – response communication and other features required for web application hosting.

![Figure 9: ContextBox components](http://www.forum.nokia.com/devices/N95_8GB/)

![Figure 9: ContextBox components](http://sourceforge.net/projects/pys60/files/)

![Figure 9: ContextBox components](http://code.google.com/intl/pl/appengine/docs/whatisgoogleappengine.html)

![Figure 9: ContextBox components](http://www.djangoproject.com/)

---

27 http://www.forum.nokia.com/devices/N95_8GB/
28 http://sourceforge.net/projects/pys60/files/
30 http://www.djangoproject.com/
Facebook role is to present the final outcome. There is no GUI\textsuperscript{31} needed, neither on the phone nor on the server site. User is not forced to take any action while the application is working. The data is collected and sent automatically to the Google App Engine storage. Figure 9 depicts components mentioned above and their communication patterns.

### 4.1.1 Mobile Context Toolbox

The original Context Toolbox application had to be customized. Some of the modules have been excluded (as it was mentioned in chapter 3.3.1 accelerometer readings, phone logs, pictures, GPS, etc. are not needed). The components that are actually used are:

- GSM
- Wireless LAN (WiFi)
- Phone information (battery state, profile)
- Bluetooth

The part of MCT which required the most customization was log synchronization module with the data storage. The initial implementation enforced communication with the CouchDB\textsuperscript{32} server using HTTP POST and JavaScript Object Notation (JSON\textsuperscript{33}) data-interchange format. This part had to be modified and adjusted to communicate with the “ConTextBox” application residing on the Google App Engine servers. On each log (raw sensor data) synchronization attempt the following steps are executed:

1. MCT context logging application (see MCT architecture diagram on Figure 5) fetches the sensor data from local data storage.
2. The sensor data is put into a JSON format (with help of simplejson Python module\textsuperscript{34}).
3. The data in JSON format is compressed with zlib Python module to reduce the amount of data to be transferred over the network (up to 60% less data to send).
4. A CRC32 checksum is calculated from the compressed data.
5. The compressed JSON data and the CRC32 checksum are merged, separated by a sign of a new line (\textbackslash n).
6. Total size of the data package is calculated in bytes and put into the “Content-length” header.
7. A POST request is created containing the data package.

---

\textsuperscript{31} Graphical User Interface
\textsuperscript{32} http://couchdb.apache.org/
\textsuperscript{33} http://www.json.org/
\textsuperscript{34} http://pypi.python.org/pypi/simplejson/
8. The package is submitted with HTTP POST to the ContextBox application using its “add” method.

9. An acknowledging response is received that contains information on whether the data package was accepted or not.

The synchronization is performed every 5 minutes under the condition that a 3G data channel is available. On each data submission the most recent 500 log entries are sent to the Google App Engine Datastore (33) storage service.

### 4.1.2 Google App Engine service

In order to reduce the amount of time required for development and configuration of a web application designed to handle the context data, Google’s App Engine service was selected. As Google claims:

> “Google App Engine lets you run your web applications on Google's infrastructure. App Engine applications are easy to build, easy to maintain, and easy to scale as your traffic and data storage needs grow. With App Engine, there are no servers to maintain: You just upload your application, and it's ready to serve your users.”(33)

The key criteria that helped to decide whether to host the application on a Unix server or on Google’s framework were the following:

- Python 2.5.x support
- Easy and speedy deployment of the web application (important for remote testing)
- Possibility to view the raw data stored in the database engine
- Public accessibility (Google provides free *.appspot.com domain)
- Minimal configuration overhead as most of the efforts were to be consumed by the development phase.

All of conditions mentioned above were fulfilled “out of the box” by Google, providing a robust hosting platform for the ContextBox application.

Along with Python 2.5 support, Google App Engine bundles in “a high-level Python Web framework that encourages rapid development and clean, pragmatic design”\(^{35}\) called Django. It leverages the Model – Viewer – Controller (MVC) architectural pattern and handles each of the layers out of the box. Google has modified Django accordingly to fit the App Engine service by providing data model and controller abstraction layers. Utilizing features and design of the

\(^{35}\) [http://www.djangoproject.com/](http://www.djangoproject.com/)
Django web development framework, the ContextBox application was build upon it and the following components were derived:

**Model**

The model represents the data upon which the application operates. In case of the ContextBox application the following models and fields are defined:

- Bluetooth
- Wifi
- PhoneInfo
- Location

Even though only the most recent context information is utilized (for the need of the profile ContextBox, see Figure 18) all of the data submitted by the mobile device is stored, as it presents a resource for potential future analysis.

The code snippet below (Code snippets 1) represents the Bluetooth model and its fields.

### Code snippets 1: Bluetooth model

```python
01 class Bluetooth(BaseModel):
02     _id = db.IntegerProperty(default=0)
03     _rev = db.StringProperty()
04     user = db.StringProperty()
05     facebook_uid = db.IntegerProperty(default=0)
06     btmac = db.StringProperty()
07     device_class = db.IntegerProperty()
08     imei = db.StringProperty()
09     name = db.StringProperty()
10     timestamp = db.StringProperty()
11     datetime = db.DateTimeProperty()
12     type = db.StringProperty(default='BTLog')
13     created_at = db.DateTimeProperty(auto_now_add=True)
```

The following fields are common for all of the models:

- `_id` – is a document ID number
- `_rev` – revision stamp of the document (single log entry)
- `user` – user name of the mobile phone holder
- `facebook_uid` – Facebook user ID of the mobile phone holder
- `timestamp` – Unix timestamp
- `datetime` – Human readable representation of the timestamp value
Translating mobile sensor data into meaningful context descriptions

- **type** – type of the log entry (BluetoothLog, WLanLog, PhoneInfoLog, LocationLog)
- **created_at** – human readable date and time value representing time at which a particular log entry was received and processed.

The code snippet below (Code snippets 2) represents a sample JSON notation of the Bluetooth model presented before. In such form the raw sensor data is transferred from the mobile device to the ContextBox application.

**Code snippets 2: Sample log data JSON notation**

```json
01 {
02   "_id": "100266578661179841",
03   "_rev": "1-d075ba55d5d4df12f55ecedab2283bb18",
04   "user": "Magda",
05   "facebook_uid": "1760582218",
06   "btmac": "00:24:03:56:4F:68",
07   "device_class": 5898764,
08   "imei": "352061025625878",
09   "name": "Milab.198",
10   "timestamp": 1269832557,
11   "datetime": "2010-02-11 21:00:25.685702",
12   "type": "BTLog",
13   "created_at": "2010-02-11 21:05:45.623470"
14 }
```

**Controller**

The controller is build by the Django team and customized by Google to suit the needs to the Google App Engine platform. Additionally Facebook development community provides Python based API\(^{36}\) which is “pluggable” into the Django middleware\(^{37}\) architecture\(^{38}\). In this case the controller is the framework itself; the machinery that sends a request to the appropriate view, according to the Django URL configuration (35).

**View**

The view is a Python callback function for a particular URL. In case of the ContextBox application there are two callback methods:

- **add** – used to accept submitted context data from a mobile device
- **locationHistory** – when called generates an image presenting a map and a history of last 10 known locations.


The “add” method is an interface for data submitted by the mobile clients. When a request comes the following process is executed:

1. The CRC32 checksum is separated from the rest of the package’s payload.
2. A CRC32 checksum is calculated from the compressed JSON payload.
3. Both checksums are compared, if the same the process execution continues. In case of a discrepancy the package is rejected and a response package containing rejection data is sent.
4. The compressed JSON data is decompressed and converted from the JSON format to a python dictionary format.
5. A 63 bit unsigned integer is generated for the purpose of a document ID.
6. Document revision is generated using the UUID module (UUID objects [universally unique identifiers] according to RFC 4122). This is done to maintain compatibility with CouchDB request – response model.
7. Type of the package is checked (Bluetooth log, WLAN log, etc.) and an adequate data storage model is called.
8. Data is stored in Google Storage.
9. JSON response is created confirming that the data package has been accepted and processed. The response contains the generated ID and document revision.
10. The latest information about the battery level, location, people around and profile setting is retrieved from the data storage.
11. A FBML\(^{39}\) narrow profile box containing the latest context information is created.
12. This FBML “message” is sent to Facebook caching servers and an appropriate user’s profile narrow box is updated. The users are matched against Facebook UID, which are embedded in every context information packet sent to ContextBox application.

The \textit{locationHistory} method utilized the Google Geolocation API\(^{40}\) which can make use of network servers to obtain a position fix. The server determines the client’s position using a set of data provided by the client (here the mobile phone). This data includes the client’s IP address and information about any cell towers (GSM) or WiFi nodes it can detect (36).

In order to obtain location information from the geolocation service, one has to POST a following JSON “token” to https://www.google.com/loc/json:

---

\(^{39}\)Facebook Markup Language, HTML-like language used to display pages inside of the Facebook canvas

\(^{40}\)http://code.google.com/intl/pl/apis/gears/geolocation_network_protocol.html
The more information the token contains the better quality result the geolocation API will return. The most significant parameters are:

- `home_mobile_country_code` - Mobile Country Code (MCC) is a number identifying the country where the mobile phone is logged on to the cell network.
- `home_mobile_network_code` - Mobile Network Code (MNC) identifies the cellular network to which the mobile device is connected to.
- `radio_type` - `wcdma` or `gsm` which is important as in some cases for the same `cell_towers` data sets different locations are returned. This might be due to separation of databases holding information about the WCDMA and GSM network cell towers.
- `cell_towers` - contains one or more records on the discovered by the mobile device cell towers. Each record set comprises of:
Implementation

- **cell_id** – carries information about the connected Base Transceiver Station (BTS) cell
- **location_area_code** – Location Area Code (LAC) which is a unique number within a network identifying location area \(^{41}\)
- **mobile_country_code** – same as the *home_mobile_country_code*
- **mobile_network_code** – same as the *home_mobile_network_code*

To raise chances of successful location retrieval and to enhance its accuracy, detected WLAN nodes’ MAC addresses are included.

Once the JSON token is submitted using the HTTP POST method to Google’s geolocation interface, a response arrives in a form as presented below:

**Code snippets 4: Sample Google geolocation response**

```json
01 { 
02   "location": 
03     { 
04       "latitude": 55.820475, 
05       "longitude": 12.497697, 
06       "address": 
07         { 
08           "country": "Denmark", 
09           "country_code": "DK", 
10           "region": "Hovedstaden", 
11           "county": "Rudersdal", 
12           "city": "Holte", 
13           "street": "Øverødvej", 
14           "postal_code": "2840" 
15         }, 
16       "accuracy": 795.0 
17     }, 
18   "access_token": "2:WkqgbYN5sBcQe7qH:5MorRKdFbMQ-wNPk"
19 }
```

According to the response specification \(^{42}\) only latitude and longitude parameters are guaranteed to be present (if the estimated location BTS location is known). The remaining fields that contain more detailed information about the location are optional and may not be included in the server’s response. The obtained location data is stored in the data storage and the most “fresh” dataset it utilized for displaying recent location in the user’s narrow profile box (ContextBox space) – see Figure 18. Further on the “last 10 known locations” are plotted again on the map and stored as an image using Google’s “Static Maps API” \(^{43}\) and static map paths. This map will be

\(^{41}\) http://www.telecomabc.com/l/lac.html  
\(^{42}\) http://code.google.com/intl/pl/apis/gears/geolocation_network_protocol.html#response  
\(^{43}\) http://code.google.com/intl/pl/apis/maps/documentation/staticmaps/
displayed in the wide ContextBox profile tab. The red marker symbolizes user’s latest location. Blue markers are previous known whereabouts, where 0 is the oldest one. Thickness of the lines joining the markers rises with the consecutive locations and denotes history of user’s estimated whereabouts. Sample data set plot on the map can be viewed on Figure 10.

![Map Image](image-url)

**Figure 10: User’s last 10 known GSM locations**

The map image presented above is recreated on each location change however it is not stored in the ContextBox application. Instead it is available via a special URL from Google Maps cache. Whenever user’s ContextBox tab is pressed, Facebook polls the ContextBox application for the latest map which is fetched from Google cache infrastructure.

The Mobile Context Toolbox has been implemented in such way that it records Cell IDs and WiFi information separately. It was necessary to find out how to connect those two logs. As the Cell ID and WiFi database records are kept separately the only link between them is the timestamp. However, as the MCT has two separate worker threads, one for WiFi logging and the second one for Cell ID logging, the timestamps for both of them differ. It was necessary to find a solution that would allow comparing those two timestamps and selecting those records that are closely matched. The implementation looks for the WiFi logs that have been saved 30 seconds before
and 30 seconds after the Cell ID log. If no records have been found the time window is expanded by another 60 seconds. The figure below (Figure 11) presents the described approach.

![Diagram](image)

**Figure 11: GSM to WiFi timestamp mapping and lookup time window expansion**

The purpose of the secondary time window expansion is to reduce the number of situations when the phone is moving fast (in a car, a bus or a bike). In that case ±30 seconds may not be enough to discover a match between the Cell ID and the WiFi networks and reducing the margin of error (WiFi location and GSM Cell location are too distant from each other).

### 4.2 Facebook platform

The Facebook Developer’s Platform (37) gives a possibility of creating a Facebook application.

![Diagram](image)

**Figure 12: Facebook application architecture. 1- Facebook server receives a URL request for an application; 2- Facebook calls the callback URL on an application; 3,4 -an application process the request getting Facebook data via API and returns FBML for display to the user; 5- Facebook process the FBML response and returns to the browser**

Facebook applications are not installed directly onto the Facebook server. The social network is responsible for the presentation of the results created by the underlying application. When the
application URL is requested, Facebook redirects the call to the server that hosts the application. After the request is processed, communication with Facebook is established via the Facebook API and FBML code is returned. Facebook transforms the FBML code into HTML and sends it into the user’s browser. General Facebook application architecture (38) is presented on Figure 12.

4.2.1 Graphical presentation of the context information
A challenging part of the implementation was visualization of the gathered data. The narrow boxes on the main profile have limited size. A usable width is 184 pixels and height is up to 250 pixels (39).

The best way to present to users clear information is to pass it on by means of icons (40) and small amount of text. Icons visually represent objects.

![Figure 13: Battery state icons](image)

The battery icon (Figure 13) illustrates the charge of the phone battery. The Python S60 `sysinfo` module offers an API for checking the system information of a S60 mobile device (by means of the MCT PhoneInfo module). The battery charge level ranges from 0 (empty) to 100 (full).

![Figure 14: Profile icons](image)  ![Figure 15: People around/Alone icons](image)

The icon, which represents a speaker (Figure 14), depicts the profile type enabled on the user’s phone. The `active_profile()` method, from Python S60 `sysinfo` module, returns the current active profile as a string, which can be one of the following: ‘general’, ‘silent’, ‘meeting’, ‘outdoor’, ‘pager’, ‘offline’, ‘drive’, or ‘user <profile value>’ (41). In configuration of silent, meeting and pager profiles the sound is muted. Again, the information mentioned above is provided by MCT.

Figure 15 presents icons that will appear depending on the user’s surroundings (showing if other phones have been detected around). A user’s geo location name is displayed next to the icon that presents the globe (Figure 16).

---

44 Uniform Resource Locator
45 Hyper Text Markup Language
As it was already mentioned, the narrow profile box has limited size. This “inconvenience” enforces a “pixel perfect” graphical design. In this case the Firebug\textsuperscript{46} Firefox extension was a real help. Firebug is the most popular web development tool for Firefox web browser. It allows the debugging, editing, and monitoring of any website’s CSS, HTML, DOM, and JavaScript. Thanks to this Firefox extension it was possible to inspect HTML and modify style and layout in real-time (Figure 17). The properly prepared box could be placed in a Facebook GUI.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{firebug.png}
\caption{Life designing with Firebug}
\end{figure}

4.2.2 Privacy settings

Facebook manages the privacy settings of the application. After installing the ContextBox on the profile, user can edit the privacy settings. This allows deciding who has access to the information that has been sent by the phone. Below are listed the available privacy settings which control visibility of the application’s Box or Tab on the profile screen:

- Everyone on Facebook
- Friends of Friends (friends and their friends can see this)
- Only Friends

\begin{footnotesize}
\textsuperscript{46} http://getfirebug.com/
\end{footnotesize}
• Some Friends (possibility of choosing specific friends who can see this)
• Only Me (only user can see this)
• Except These People (possibility of choosing specific friends who cannot see this)
5 Evaluation

In this chapter the tests of the created application will be described. The functionality and noticed issues will be illustrated.

Additionally, a survey that examines the general users' opinion on Facebook and possibility of sharing context will be presented.

5.1 ContextBox evaluation

The system has been tested by 2 persons for the duration of 10 days. The users, after installing the ContextBox service on their Facebook accounts, were equipped with the Nokia N95 smartphone with customized Mobile Context Toolbox (MCT) application.

The first and the only step, to start synchronizing the phone logs with Facebook account, is to enter the Facebook user ID in the MCT settings.

The application is “comfortable” and not demanding to use. The only action that is required from the user is to input the Facebook user ID and then the application will synchronize automatically with the correct Facebook account.

The narrow box (Figure 18), on the left side of Facebook profile page, has been updated accordingly to the data sent by the phone. This information constitutes the general picture of user availability.

![Figure 18: Processed context information (narrow ContextBox) on the user's profile page](image)
During the international travel the MCT application “crashed”, which was probably caused by the Python data base storage service. From observations and testing it seems that the issue is not only connected to the number of SQL queries executed and the size of the database file. As the application crashed at random times, sometimes when the database file size was below 1MB and the number of queries did not go over 2039 (as suggested on the Forum Nokia discussion board).

The problem seems to be fixed by reloading the “e32db” module at certain intervals of time, but this solution requires more testing to be considered as stable.

The map is “clear” only if the phone moves constantly and does not return to the same place in short intervals of time. In situations when the user is walking slowly back and forward between two different Cell IDs, the lines on the map can overlay each other creating an unreadable blob of colors.

The biggest problem with the acquiring the location is that without a GPS it is very hard to determine the exact whereabouts. The point on the map can only suggest the area in which user is located. However the GPS usage is pointless in this case because the task of the application is to run uninterruptedly for long time and GPS utilization has a negative impact on the battery life. Another issue is that sometimes the Google maps infrastructure could not find the detected Cell ID and did not return the proper coordinates. That caused impossibility of determining user’s current location. Additionally, the map shown on Figure 10 does not reflect the exact user’s locations. The marked points have been determined only basing on Cell ID information (there was no data containing WiFi positions). WiFi location is strictly connected with the “Google Street View” service as Google’s car is collecting the following data: photos, 3-D building imagery and the most importantly WiFi network information. All of that data is correlated with the GPS coordinates.

A potential bottleneck is a situation when for a long time there is no possibility to synchronize the data with Facebook. All the logs are stored locally on the phone. The solution could be based on deleting the oldest information or compressing it to avoid running out of phone storage. The storage consumption is not the only disadvantage caused by the lack of the communication with the backend service. This problem can also bring some discrepancies between user’s actual status and the data presented on the Facebook profile. The information might be highly outdated.

---

Future work

In the future the application’s usability may be extended by bringing back the Mobile Context Toolbox’s modules that have been excluded. Additionally, a detection of nearby people’s identities can be added (e.g. by comparing the Facebook user ID with the phone Bluetooth MAC address). Also, expanded privacy setting will be needed. It would be a good idea if users could change those settings separately for each of the pieces of the shared information (phone info, location, etc.).

5.2 Survey

In order to get to know the Facebook users’ opinion about the service and the possibility of sharing context information on a social network site a small survey has been conducted. The questionnaire has been filled out by more than 60 people (mainly students, aged 25-30). Only 4 of the asked persons claimed that they do not have a Facebook account. The reason that has been given is “Not interested”.

The surveyed Facebook members answered a list of questions that have been divided into 5 categories. The aims of each questionnaire’s section are listed below:

- assess satisfaction level of Facebook in general
- assess Facebook usage scenarios
- assess privacy specific settings of Facebook
- assess common Facebook activities
- assess an opinion on sharing context information on Facebook

The survey has been created on Google Documents service and is available on http://bit.ly/9ediuw. The complete list of questions and answers is presented in the Appendix B.

5.2.1 General satisfaction level of Facebook

More than 70% of surveyed persons have a Facebook account for 1,5 – 2,5 years. Most of them are neutrally satisfied with the service. Only 14% think that this social network really improve their life (Figure 19). A slender amount (5%) of users is not satisfied at all with the features that Facebook offers. But despite the fact that they are disappointed, they still have an active Facebook account.
5.2.2 Facebook usage scenarios

Most of the surveyed people (51%) dedicate less than 10 minutes a day on using Facebook (Figure 20). Only 2 persons (3%) declare that they spend more than 3 hours a day monitoring the website.

The question “if Facebook is a part of someone’s everyday activity” does not have a straight answer (Figure 21). Almost the same amount of people claim that they “Agree” (29%) and that they “Strongly disagree” (31%).
Additionally, the majority (42%) believe that they do not feel out of touch when they have not logged on to the Facebook for a while.

After analyzing those results it is possible to come to a conclusion that Facebook is “a place” where people like to check in from time to time, but it is not too absorbing. Of course there are individuals that spend more significant amount of time on Facebook (more than 3 hours a day). The service is used mainly for communication with friends (the current one and the old one).

5.2.3 Privacy specific settings of Facebook

It seems that most of the Facebook members (75%) are using the service consciously. They are aware of who has access to the presented information.

![Privacy settings](image1)

**Figure 22: Privacy settings**

Unfortunately, still 7% of the surveyed people do not know who can see their Facebook profiles (Figure 22) and some of them accept “friend invitations” from people they do not know.

![Providing email address and phone number](image2)

**Figure 23: Providing email address and phone number on Facebook**

Another observation (Figure 23) is that users more often present on Facebook their email address (73%) then the phone number (12%). The phone number information seems to be more confidential, even if the access to the profile is granted only to the people from the friends list. It probably results from the thinking that it is easier to ignore an unwanted email than the ringing phone.
5.2.4 Common Facebook activities
The average amount of the declared Facebook friends is 100 people. Of course this number is growing. Members are meeting new people in the real life and then invite each other to the social network’s site friend’s list.

One of the most popular activities on Facebook is photos sharing.

5.2.5 Sharing context information on Facebook
More than 60% of the surveyed persons are still not interested in sharing context information on social network sites (Figure 24). They do not want other people to know where they are and what they are doing.

![Figure 24: Willingness to share context information on Facebook](image)

From the 37% of users, who do not have anything against uploading additional information to the network, a significant majority (82%) would like it to be done automatically (Figure 25). Manual “tagging” is not a desired way of collecting data.

![Figure 25: The way of collecting sensor data](image)

From the list of possible pieces of information (Figure 26) the calendar events has the fewest amount of votes (36%). The current location (82%) is the most wanted record to be displayed on Facebook. The preferable way of showing the whereabouts information (Figure 27) is to determine an approximate area (for example a name of town or a region within a chosen
radius). Additionally, a possibility to determine a group of people, who can see more detailed information, would be welcome. Another noteworthy observation is that over 60% of surveyed Facebook users are **not interested in sharing the locations history** (Figure 28).

The **“decide per group”** is also the most preferable option for people who would like to share their activity (Figure 29) and calendar events (Figure 31). Almost the same results can be seen on
Figure 30 which shows that users would like to decide who has access to the information about their companionship. Facebook users may as well just inform others that they are not alone.

**Figure 29: Willingness to share activity information**

**Figure 30: Willingness to share “people around” information**

**Figure 31: Willingness to share calendar events information**

**Figure 32: Willingness to see friends’ context information**
The survey shows that 18% of all the users, who would like to share context, are not interested in seeing the same information on their friends’ profiles (Figure 32). However, a noteworthy observation is that higher amount of users want to observe data presented by their friends, than post it by themselves (Figure 26 and Figure 33).

5.2.6 Summary

The survey was submitted by more than 60 participants. Those were people in the age between 24 and 30, mostly students. Most of the surveyed individuals think that they are neutrally satisfied with the features that Facebook offers to them. A slender amount of the network users declare that they spend more than an hour a day monitoring the website, the majority dedicate only 10 minutes a day. The network is used mainly with the aim of keeping in touch with friends and only the known people have access to the Facebook profile. Unfortunately, still some of the network users do not care about their privacy settings (they do not know who has access to their profiles, photos, personal information). It can be caused by ignorance or just by not being aware of the possibility of changing the privacy settings.

The most interesting are the results concerning the willingness to present personal context information. The statistics show that majority does not want to share too much personal information about their life with others. Out of the group, who would like to post their context on Facebook, most of the people are looking for the solution that would allow them to control their privacy settings in a detailed fashion (who exactly has access to what information). For example, they only share the location with people they are comfortable meeting up with, and
when they want to be found. The least amount of votes was given to the possibility of sharing the phone’s calendar events. Probably this information is considered as too private. There is a huge gap between telling that “I’m busy” and announcing “visit at neurologist at 10”.

Finally, one of the humans’ characteristic features has to be mentioned – curiosity. Facebook is popular not only because users like to share their personal activities, but also because they want to see others people life.
6 Discussion and Conclusions

The aim of this thesis was to analyze the topic of extracting user’s context information from the sensor data collected by a mobile phone. Additional goal was to design and implement a prototype of an application that presents this context on a chosen social networking site. This chapter is dedicated to observations made while working on this thesis.

Mobile phones transform from simple “selective” communication devices to versatile tools. Users can communicate with each other not only by calling and sending short messages (SMS), but also by posting information about their state, even without touching the phone. This possibility is offered by sensors build in almost all the new generation mobile phones.

Sensors have begun to infiltrate people’s everyday life. As it has been described in chapter 2, the collected individual user’s context data can result in getting to know his daily schedule. With the help of a mobile phone (GPS, GSM logs), it is possible to find, among other things, user’s location. By analyzing this information and combining it with other data (date and time, duration of stay in one location, etc.) it is probable to detect meaningful places (e.g. home, work). For some people this information may be an extremely interesting research topic. For others those are just useless statistics and some groups can even be terrified of being under constant surveillance.

Sensor data can be a perfect source of information about human’s availability. By selecting the “silent” phone’s profile, user gives a straight message that says: “do not disturb me”. Unfortunately, the only person who knows the profile settings is the phone’s owner himself. Other people, who are trying to contact him, know only that no one is picking up the phone.

As Internet becomes a natural communication tool, the attractive solution seems to be the possibility of sharing the context (especially the availability information) on a social network. When people are sitting in front of the monitor, they do not call somebody in order to ask a question. The first thing that is done is checking if the needed person is active on an Internet messenger (for example the so popular Gmail chat). Another phenomenon is that sometimes in order to send an email to somebody, people have to do it through the social network site. The reason is that their friend does not check the standard mailbox every day, but systematically controls the social network account.

Here comes Facebook’s role, a social network site that is already in demand. Facebook has been chosen as an environment for the practical part of this thesis. As the user’s status message on this network is very often used like Tweeter (members announce where they are or what they
are doing) a good idea was to systematize this information. Other deciding factors, which helped during the selection process of a social network, were the Facebook popularity and the fact that very often “Facebook friends” are also mates in the real life. This eliminates a bit the privacy problem (if user wants to show his context, he can control who can see it).

In a social network’s world, the success of the system is tied to the amount of contribution that members have produced. This is the reason, why it is so easy to find the network that allows developers to create new applications (in this case Facebook). Additionally, the Mobile Context Toolbox (MCT) and the Google App Engine platform represent a flexible environment for context data collection, storage, processing and presentation. By joining those three structures (MCT, Google service and Facebook) it is possible to create a new tool that can enrich the social network user’s profile page. The profile can not only bring enjoyment to Facebook member’s friends but also may be useful (show availability of the user).

The recipient of the created application is not the user himself, but all the friends that would like to know his basic status information (availability, location etc.). Sometimes very detailed information is not needed. An example can represent the location. User’s mates usually know his address, and the approximate area (with the name of the district or the street) is for them enough to find out if the friend’s location is for example home. This assumption can be supported by the survey results. Only 1 person gave the straight answer that he/she would like to present the exact point on the map as his/her whereabouts.

The visualization and sharing of the data are not the only problems which need to be overcome. The biggest challenges can be seen in terms of how to approach the data (analysis) and understand it. The more facts are known about the user, the more useful are records from the sensors. In general, phone owner has to “help” the application to collect as much information as possible (one of the actions is to carry the phone for almost all the time). The fastest way, of learning the user’s behaviors and habits, can be gained by tagging the meaningful location and activities so the system can get to know the correlations between them (e.g. @work = working/busy). Unfortunately the tagging procedure is pointless if the assumption of the application is to be completely transparent for the user.

Of course the phone can “listen and learn” by itself. Unfortunately, if there is only a short period of time allocated for the data collection, it is almost not possible to achieve the last level of the context “pyramid” (Figure 1). Even some of the “knowledge” elements may be not enough precise. An example can be a situation when the user starts the application during his vacation
and the system assigns a hotel location to the “@home” label. For a long initial period of time the data analysis in part will be founded on probabilistic speculation.

The idea of the application was to design a prototype system that does not need any user’s attention. The data collection and synchronization with the social network should be done automatically. This type of design is more convenient for the user but at the same time more complicated to implement. In other solutions, where user is a part of the system (has to tag places, decide when post the data, etc.) it is easier to find problems with information analysis and technical glitches of the system. While the phone is hidden in a pocket or bag it is not possible to immediately see when the application is not working properly.

The implemented system is presenting users’ availability and their approximate location. In the future MCT could be easily further customized. The modules that have been excluded in the implementation may be brought back in order to extend the application’s usability and help achieve the “wisdom” level of the context pyramid (Figure 1).

Realization of a system that listens, learns and returns translated and usable user’s context is a great achievement that is yet to come. It is a huge challenge to reach the highest context level (“wisdom”) without direct user’s help. All the research and analysis in the world may be insufficient in situations when the phone is left behind and does not reflect user’s current status. Human race and the technology have to be more inseparable. However, users are individuals with emotions and the way they behave is related to their current mood. It means that even an emotional state does impact the willingness to disclose somebody’s shared context (e.g. location). For example, when a user is stressed or angry, he might not want anyone to know where he is (42). Another matter constitutes gaining users’ trust so they will willingly and readily share their context data. It is connected to the non trivial privacy problem which needs to be addressed and handled properly for each of the participating users.

A separate matter represents broader and more common access to the “mobile” Internet as still in some of the locations in Europe there are blank spots on the map of cellular coverage. Another issue is to have a well described (metadata) cellular infrastructure so it will be possible to map GSM cell IDs and correlate them with WLAN locations to narrow down the location radius. Tests of the ContextBox application show that there still is a huge gap between different countries when it comes to GSM metadata.

The topic of translation of the mobile sensor data into meaningful context descriptions is an open research topic. There are still new, different ideas and methodologies of contextual data
analysis and shaping to explore. The results may represent a system that enables a better information transition to knowledge, and from knowledge to wisdom.
References


33. **Google code, Google App Engine.**

34. **Sanderson, Dan.** Programming Google App Engine Build and Run Scalable Web Apps on Google’s Infrastructure. *O’Reilly Media.* November 2009.

35. **Django MVC description, Website.**

36. **Google code, Geolocation API Network Protocol, Website.**


39. **Facebook Developers, New Design Narrow Boxes.**


41. **Nokia.** PyS60 Library Reference Release 1.4.5 . 3 December 2008.


43. **Dye, Mark.** Lifestreaming and the digital you. 24 Nov 2009.


45. **Dosanjh, Ash.** Mobile web users head to social network sites.

46. **WORLD have your say, Website.**
Appendix A

Facebook profile site view with marked ContextBox area.
Appendix B

Survey - the complete list of questions and answers.

64 responses

Summary See complete responses

A) Do you currently have a Facebook account?

- Yes: 60 (94%)
- No: 4 (6%)

B) If no, why not?

- Do not have regular computer access: 0 (0%)
- Do not have the time: 1 (25%)
- Not interested: 4 (100%)
- Have never heard of Facebook before: 0 (0%)
- Other: 0 (0%)

People may select more than one checkbox, so percentages may add up to more than 100%.

If you answered “Yes” to question A, please continue. If “No”, please submit the questionnaire by pressing the “Submit” button on the bottom of the email. Thank you for your time.

The first part of this questionnaire is to assess more general attitudes of Facebook.

Approximately how long have you had your Facebook profile? - Years

- 1 or less: 7 (12%)
- 1.5: 16 (27%)
- 2: 20 (34%)
- 2.5: 11 (19%)
- 3+: 5 (8%)

The percentages do not add up to 100% because respondents could select ‘Other’.
Appendices

How satisfied are you with Facebook, overall?

- Not satisfied at all: 3 (5%)
- Barely satisfied: 10 (17%)
- Neutral Satisfied: 38 (64%)
- Very satisfied: 8 (14%)

Which of the following best describes the Facebook application from your perspective?

- It is how I communicate with my current friends: 27 (47%)
- It provides a distraction from my schoolwork: 11 (19%)
- It allows me to communicate with people from my past: 29 (50%)
- It allows me to collect information on people I am interested in: 22 (38%)
- It provides me with information (e.g., in groups): 13 (22%)
- Other: 8 (14%)

People may select more than one checkbox, so percentages may add up to more than 100%.

This section of the questionnaire is to assess your basic use and attitudes towards Facebook.

On average, approximately much time per day do you spend on Facebook?

- 10 min or less: 30 (51%)
- 10-30 min: 9 (15%)
- 30min – 1h: 11 (19%)
- 1-2 h: 7 (12%)
- 2-3 h: 0 (0%)
- 3+ h: 2 (3%)

Facebook is part of my everyday activity

- Strongly Disagree: 17 (29%)
- Disagree: 8 (14%)
- Neutral: 12 (20%)
- Agree: 18 (31%)
- Strongly Agree: 4 (7%)
I dedicate a part of my daily schedule to Facebook

<table>
<thead>
<tr>
<th>Rating</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>19</td>
<td>32%</td>
</tr>
<tr>
<td>Disagree</td>
<td>19</td>
<td>32%</td>
</tr>
<tr>
<td>Neutral</td>
<td>13</td>
<td>22%</td>
</tr>
<tr>
<td>Agree</td>
<td>8</td>
<td>14%</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

I feel out of touch when I haven't logged on to Facebook for a while

<table>
<thead>
<tr>
<th>Rating</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>25</td>
<td>42%</td>
</tr>
<tr>
<td>Disagree</td>
<td>16</td>
<td>27%</td>
</tr>
<tr>
<td>Neutral</td>
<td>11</td>
<td>19%</td>
</tr>
<tr>
<td>Agree</td>
<td>5</td>
<td>8%</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>2</td>
<td>3%</td>
</tr>
</tbody>
</table>

This section of the questionnaire is to assess your privacy specific settings of Facebook

Who can see your Facebook profile?

<table>
<thead>
<tr>
<th>Privacy Setting</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Everyone</td>
<td>8</td>
<td>14%</td>
</tr>
<tr>
<td>Only my friends</td>
<td>44</td>
<td>75%</td>
</tr>
<tr>
<td>Don't know</td>
<td>7</td>
<td>12%</td>
</tr>
<tr>
<td>Only my friends [44]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Who is allowed to see the following features on your Facebook profile? - Status Updates

<table>
<thead>
<tr>
<th>Privacy Setting</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only my friends</td>
<td>44</td>
<td>75%</td>
</tr>
<tr>
<td>Everyone</td>
<td>5</td>
<td>8%</td>
</tr>
<tr>
<td>Don't know</td>
<td>10</td>
<td>17%</td>
</tr>
</tbody>
</table>

Who is allowed to see the following features on your Facebook profile? - Photos tagged of you

<table>
<thead>
<tr>
<th>Privacy Setting</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only my friends</td>
<td>46</td>
<td>78%</td>
</tr>
<tr>
<td>Everyone</td>
<td>5</td>
<td>8%</td>
</tr>
<tr>
<td>Don't know</td>
<td>8</td>
<td>14%</td>
</tr>
</tbody>
</table>
Appendices

Who is allowed to see the following features on your Facebook profile? - Online Status

- Only my friends: 45 (76%)
- Everyone: 4 (7%)
- Don't know: 10 (17%)

Who is allowed to see the following features on your Facebook profile? - Wall

- Only my friends: 44 (75%)
- Everyone: 5 (8%)
- Don't know: 10 (17%)

Do you use the “Block List” to prevent certain people from searching for you?

- Yes: 11 (19%)
- No: 44 (75%)
- Don’t Know: 4 (7%)

If yes, why?

- To avoid certain people whom I do not want to communicate with: 8 (73%)
- To prevent certain people from “stalking” me: 7 (64%)
- Other: 1 (9%)

People may select more than one checkbox, so percentages may add up to more than 100%.

Do you use the “Limited Profile List” to prevent certain people from seeing certain aspects of your profile?

- Yes: 15 (25%)
- No: 37 (63%)
- Don’t Know: 7 (12%)
This section of the questionnaire is to assess your use of the basic Facebook functions and applications.

### How many friends are on your Facebook Friends List?

- 10 or less: 43 (73%)
- 10 - 30: 6 (10%)
- 30 - 60: 10 (17%)
- 60 - 100: 18 (31%)
- 100 +: 25 (42%)

### Approximately how many Photo Albums do you presently have on Facebook?

- 5 or less: 42 (71%)
- 5 - 10: 14 (24%)
- 10 +: 3 (5%)
Appendices

What/who is on your pictures?

- Family: 19 (34%)
- Significant Other: 12 (21%)
- Friends: 38 (68%)
- Pets: 9 (16%)
- Parties: 27 (48%)
- Myself: 42 (75%)
- Scenery: 16 (29%)
- Sporting Events: 5 (9%)
- Art: 2 (4%)
- Other: 2 (4%)

People may select more than one checkbox, so percentages may add up to more than 100%.

C) If you had possibility to upload to your Facebook profile page data collected by you phone (e.g. current status, activity, location), would you like to do it?

- Yes: 22 (37%)
- No: 37 (63%)

D) If no, why not?

- Do not want people to know where I am and what I'm doing: 29 (94%)
- Other: 3 (10%)

People may select more than one checkbox, so percentages may add up to more than 100%.

If you answered “Yes” to question C, please continue. If “No”, please submit the questionnaire by pressing the “Submit” button on the bottom of the email. Thank you for your time.

How would you like the data to be collected?

- Manual input (e.g. by tagging place @home@work): 5 (23%)
- Automatically by ...: 18 (82%)

People may select more than one checkbox, so percentages may add up to more than 100%.
Translating mobile sensor data into meaningful context descriptions

If you selected YES for Location, how detailed the information should be?

- Exact address (point on the map): 1 (5%)
- Approximate area (e.g., town): 7 (37%)
- General name of the place (home, work, school): 4 (21%)
- Decide per group (family, friends, etc.): 6 (32%)

Would you like to show history of locations (e.g., 10 last approximate locations)?

- Yes: 7 (35%)
- No: 13 (65%)

If you selected YES for Activity, how detailed the information should be?

- Exact activity (e.g., walking, going by car, going by bike): 3 (25%)
- General info (moving, stand still): 1 (8%)
- Decide per group (family, friends, etc.): 7 (58%)

If you selected YES for the People around, how detailed the information should be?

- Just inform that I am not alone: 6 (43%)
- Give names of the people around: 2 (14%)
- Decide per group (family, friends, etc.): 5 (36%)

Detailed event information (in a meeting, at work, school, etc.)
- Decide per group (family, friends, etc.): 2 (25%)
- General information (availability information): 3 (38%)
- Decide per group (family, friends, etc.): 3 (38%)
Appendices

Would you like to see this information on your friends' Facebook profile?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
<td>4</td>
</tr>
</tbody>
</table>

If YES what type of information?

<table>
<thead>
<tr>
<th>Information Type</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>20</td>
<td>100%</td>
</tr>
<tr>
<td>Activity</td>
<td>12</td>
<td>60%</td>
</tr>
<tr>
<td>Profile of the phone</td>
<td>14</td>
<td>70%</td>
</tr>
<tr>
<td>Battery state</td>
<td>11</td>
<td>55%</td>
</tr>
<tr>
<td>People around you</td>
<td>12</td>
<td>60%</td>
</tr>
<tr>
<td>Calendar events</td>
<td>8</td>
<td>40%</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>

People may select more than one checkbox, so percentages may add up to more than 100%.

If NO, why?

<table>
<thead>
<tr>
<th>Reason</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not interested</td>
<td>4</td>
<td>80%</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>20%</td>
</tr>
</tbody>
</table>

People may select more than one checkbox, so percentages may add up to more than 100%.