Jun Wu

Signal Collecting Platform and “Handprint” Positioning System

Master’s Thesis
Espoo, June 30, 2014

Supervisors: Professor Antti Ylä-Jääski, Aalto University
Professor Jeong-woo Cho, The Royal Institute of Technology, Sweden
Instructor: Dr. Zhonghong Ou
Mobile computing is an emerging research field in recent years. Although the computation capability of mainstream smartphones are several orders of magnitude better than computers twenty years ago, the capacity of battery does not increase at same speed. To save energy, some recent work tries to schedule network traffic according to signal strength variations. To achieve this goal, a database that is used for storing signal strength distribution is essential.

We first design and implement a platform to collect cellular network information, including Cell-ID and signal strength information. The platform is designed as a distributed system that supports collecting signal strength data by using crowdsourcing approach.

We then deploy the platform and collect signal strength information in Otaniemi area (Finland). After analysing the collected data, we observe several interesting phenomena. (1) the density of base stations is out of expectation; (2) cells is becoming smaller; (3) in most places a device may connect to different base stations. Based on these observations, we design a new energy-efficient positioning system called “handprint”, which utilises signal strength information from neighbouring smartphones to assist positioning. Compared with Google Geolocation API and other existing work, our “handprint” system can improve positioning accuracy by more than 20%.

**Keywords:** crowdsourcing, positioning, cellular network, short range communication
The process of completing this master thesis is exciting and wonderful. I am proud of this work.

Thanks my supervisor professor Antti Ylā-Jääski, professor Jeong-woo Cho, and my instructor Zhonghong Ou. Without their valuable suggestions and instructions, I can’t believe I can complete such a challenging job.

Thanks NordSecMob program, which brings me with such a memorable experience in Nordic countries.

Thanks everyone who has helped me in the past two years. In return for your kindness, I will always click “likes” on your Facebook pages.

Finally I would like to thank deadline, it is always the drive force to push me work hard.

Espoo, June 30, 2014

Jun Wu
Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wi-Fi</td>
<td>Local area wireless technology based on IEEE 802.11 standards.</td>
</tr>
<tr>
<td>NFC</td>
<td>Near Field Communication</td>
</tr>
<tr>
<td>LTE</td>
<td>Long-Term Evolution, 4th generation mobile communication standard</td>
</tr>
<tr>
<td>UMTS</td>
<td>3rd generation mobile cellular system for networks based on the GSM standard.</td>
</tr>
<tr>
<td>CDMA2000</td>
<td>A family of 3rd generation mobile technology standards.</td>
</tr>
<tr>
<td>LBS</td>
<td>Location Based Service</td>
</tr>
<tr>
<td>CAPS</td>
<td>An energy-efficient positioning for smartphones using cell-id sequence matching.</td>
</tr>
<tr>
<td>WPAN</td>
<td>Wireless Personal Area Network</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Position System</td>
</tr>
</tbody>
</table>
# Contents

Abbreviations and Acronyms  

1 Introduction  
1.1 Problem statement  
1.2 Structure of the Thesis  

2 Background  
2.1 Platform  
2.2 Positioning system  
2.2.1 Satellite based positioning system  
2.2.2 Triangle algorithm based positioning system  
2.2.3 Cell-ID aided positioning system  

3 Signal collecting platform  
3.1 Architecture  
3.2 Database  
3.2.1 Data compression  
3.2.2 Difference between models  
3.3 Middleware  
3.4 Client-side  
3.4.1 Signal collecting  
3.4.2 Helsinki Bus Timetable  
3.4.3 Research app  
3.5 Extension  

4 Observations  
4.1 Density of base stations  
4.2 Small cells  
4.3 Intersection of multiple base stations  
4.4 Further observation of multiple coverage
5 Handprint positioning system
   5.1 Background ............................................. 37
   5.2 Handprint ............................................. 38
   5.3 Implementation details ................................. 39
      5.3.1 How to share fingerprint information ............ 39
      5.3.2 How to estimate position ........................ 42
   5.4 Performance evaluation ................................. 42
      5.4.1 Experiment scenario ............................... 43
      5.4.2 Compare with CAPS ............................... 44
      5.4.3 Compare with Google Geolocation API ............ 44
   5.5 Converging speed ..................................... 45
   5.6 Power consumption .................................... 48

6 Discussion and future work ............................... 49
   6.1 Knowledge Share Protocol ............................. 49
   6.2 Future improvements on the prototype ............... 51

7 Conclusions ............................................. 52
Chapter 1

Introduction

Mobile computing has been experiencing a revolution in the past few years. In 2007, Apple Inc. released its new product, named iPhone, a phone without keyboard. With the great success of iPhone, smartphone rapidly became the indispensable companion in people’s everyday life. A smartphone usually consists of the following components: a multi-touch screen, various sensors (including camera, accelerometer, magnetic sensor and gyroscope), short range communication modules (such as Bluetooth [5], Wi-Fi and Near Field Communication [14]), advanced chips whose computing power is equivalent to computers several years ago, a smart operating system that supports hundreds of applications.

Let’s recall the following moments in daily life. When you get up in the morning, the first thing is to check the local weather forecast and bus timetable. If you have a car, you can also let the phone choose the route to your working place according to traffic conditions. To save time, you can use the phone’s embedded NFC function for payment when boarding a bus. From this moment on, the phone is your entertainment hub in the boring journey. You can read news and books, receive and send emails, watch videos from YouTube, play games and click “like” on friends’ Facebook. At lunch time, deciding where to eat is usually difficult but now there are many location based apps that will recommend nearby restaurants sorted by their ratings. Similar services are offered for other kinds of commercial entities such as shops, cinemas, barbershop and gyms. With smartphone, users are connected to Internet 24 hours a day, and seven days a week.

Recently Cisco published a white paper for 2013 global mobile data traffic [9]. Global mobile data traffic grew 81% in 2013 and reached 1.5 exabytes per month by the end of 2013. This is approximately 18 times the size of the overall global Internet in 2000. 88% of the traffic is generated by smart devices. Figure 1.1 projects the traffic increase until 2018. To
meet the explosive increasing requirements of mobile users, cellular network is frequently upgraded. The fourth generation technology, i.e., Long Term Evolution (LTE), can offer download speed up to 100 Mbit/s and upload speed up to 50 Mbit/s. However, currently the third generation technology, i.e., Universal Mobile Telecommunications System (UMTS) and CDMA2000, is still the mainstream technology that has been widely deployed used. Furthermore, the second generation technology, e.g., GSM, is still serving and will serve a large number of mobile users for a while.

The explosive increase of device, infrastructure and data traffic also brings many new challenges. One of the major concerns is energy consumption. Compared with the increasing of device’s power consumption, the capacity of battery lags behind in recent years. Nowadays, mobile users have to charge their smartphones every day while ten years ago the charging frequency is once every several weeks. Due to this reason, academia and industry have improved the state-of-the-art covering various aspects: improving the transmission efficiency for wireless interfaces, adopting better policy to save energy for Wi-Fi sensing [19], designing smart interaction way and ingenious algo-

Figure 1.1: Traffic increasing trend [9]
algorithms to reduce the working time of screen, CPU and sensors, inventing new material and power saving components.

1.1 Problem statement

According to a recent study by Carroll and G Heiser [7], wireless communication interfaces, especially the interface of cellular network, contributes a prominent part in the overall power consumption. Moreover, the power consumption of wireless component is not constant all the time. It is intuitive that the transmission consumes more power when the signal strength is bad. Schulman et al. [31] proved this relationship and developed an energy-aware traffic scheduling system called “Bartendr”. Bartendr divides workloads into two types: syncing and streaming. Corresponding to their features, Schulman et al. [31] chose a threshold-based technique and used dynamic programming for optimal scheduling.

Bartendr [31] demonstrated the possibility of intelligent scheduling by predicting signal strength. Nevertheless, it did not provide a complete solution for practical deployment. One major challenge that impedes the idea to become practical is the way to collect signal strength. Existing signal propagation models are based on unrealistic assumption that the space between the mobile device and the base station is vacuum. In reality, the world is much more complex. For example, obstructions such as high buildings and trees introduce serious bias for vacuum-based models. To make the system work, a signal strength map has to be generated and maintained in real time.

Motivated by Bartendr, Ou et al. [26] propose to use crowd-sourcing approach to solve this problem. They conducted a systematic study to explore the relationship between the power consumption of smartphones and all possible factors, including weather, season and phone’s model. Based on these observations, they proposed a smoothing algorithm to process signal strength, defined the concept of traces and created a series of rules for implementation. Besides that, they evaluated the traffic scheduling mechanism on selected traces in Finland with their own algorithms. The prototype achieved 18%-35% power saving, depending on the application categories.

Luckily, as a research assistant, I participated the work presented in [26] in 2013, and this thesis is a continuation and extension of [26]. The first half of the thesis will focus on building a platform to collect information about cellular network using crowd-sourcing approach. With this platform, we collect signal information in Otaniemi area (Finland). After analysing the dataset, we observe several interesting phenomena and propose a novel Cell-ID based positioning system.
CHAPTER 1. INTRODUCTION

The second half of the thesis deals with designing and implementing a new Cell-ID based energy-efficient positioning system named “Handprint”, which relies on the platform mentioned above. Compared with other existing positioning solutions, the “Handprint” system can achieve good accuracy without extra infrastructure, e.g., GPS. Furthermore, “Handprint” is not impacted by complex obstructions in the surrounding environment.

1.2 Structure of the Thesis

The thesis consists of seven chapters. Chapter 1 and Chapter 2 introduce the background and motivation of this thesis. Chapter 3 briefly describes the design and implementation of the crowdsourcing based platform. Based on data collected by the platform, Chapter 4 discusses interesting phenomena observed in cellular network. Chapter 5 proposes an innovative positioning system named “Handprint”. Chapter 6 discusses future work and Chapter 7 concludes the thesis.
Chapter 2

Background

2.1 Platform

Several crowd-sourcing based projects has been existing. OpenSignal \(^1\) is a project aiming to build a wireless coverage map with the data collected from the users of OpenSignal application. Users not only contribute to the project, but also benefit from it, which is the core principle of crowd-sourcing concept. For example, one of the OpenSignal apps can help users find nearby Wi-Fi access point. Another interesting service of the platform is to release network coverage ranking report periodically. Users can compare the coverage area and average signal strength among different operators. Specifically, it can offer a map that uses heat-map to represent signal strength. This service provides an efficient way for users to choose cellular network operator and data plan. In 2013, OpenSignal was named the “UK’s Most Innovative Mobile Company” by UK Trade & Investment [1].

Telematics company ENAiKOON launched another similar project OpenCellID \(^2\). It is a collaborative community project whose purpose is to collect GPS coordinates of cell towers. OpenCellID also provides real time statistics for cell tower distribution and a heat map of users.

However, neither of these two platforms can satisfy the requirements of energy saving scheduling systems mentioned above. OpenCellID is cell tower oriented. In other words, there is no information store/query mechanisms for arbitrary position. Furthermore, signal strength is not a mandatory item in OpenCellID’s dataset. Although OpenSignal provides an API to query the network ranking in a specific area, the granularity is as high as several kilometres. The limitation of request frequency is another limiting factor for

\(^1\)http://http://opensignal.com/
\(^2\)http://opencellid.org/
2.2 Positioning system

As a research topic with relatively long history, positioning has been existing for a while, and various positioning systems have been proposed for different scenarios. There are multiple ways to categorise them. From the perspective of usage scenarios, they can be classified as indoor versus outdoor. Indoor positioning system usually exploits Wi-Fi access points and sensor networks, which is largely different with outdoor positioning system.

Base on the underlying techniques, outdoor positioning system can be divided to three categories, i.e., satellite based, triangle based, and Cell-ID aided positioning systems.

2.2.1 Satellite based positioning system

The first category of outdoor positioning system is satellite based. Users of this system need a specialized radio receiver to receive, recognize and analyse the signals acquired from satellites. Thus, satellite based systems can cover most of surface area of Earth, as well as near-Earth space.

Existing and planned satellite based positioning systems include: Global Positioning System (GPS), GLONASS [23], Galileo [10], Beidou [37] navigation system and etc.

Global Positioning System project is created by U.S. Department of Defense (DoD) from 1973 and fully operational from 1995 [16]. As the first global outdoor positioning system, it was soon applied to civil industry. With a GPS receiver, a user will obtain his longitude and latitude after acquiring information from more than four satellites. The accuracy of GPS varies from several meters to tens of meters. Good accuracy and coverage area make GPS receiver widely integrated to aero-crafts, navigators, mobile phones, tablets and wearable devices.

Nevertheless, weaknesses of GPS are also obvious. First, users have to wait for 10 to 30 seconds before they could acquire the coordinates. This time is used to acquire information from enough satellites. If unfortunately the environment is complex, e.g., there are a lot of obstructions between the device and satellites, the initial searching process will be even longer. Second, the energy consumption of GPS becomes a heavy burden for mobile and wearable devices due to their limited battery capacity. Many measurements showed that the power of enabled GPS is more than 300 mW for smartphones [12] [11]. This high power-consumption makes GPS the most...
power hungry component besides screen and wireless communication modules. Moreover, this situation will become even more serious with the popularity of location based services [20] and applications, because GPS module is turned on for ever-increasing amount of time.

2.2.2 Triangle algorithm based positioning system

The second category of outdoor positioning system is triangle algorithms based positioning system. Different from satellite based solutions, these methods estimate the distance between the device and three nearby base stations and calculate positions using triangle algorithms. According to the metric of distance estimate, they can also be classified as time of arrival (ToA) method, angle of arrival (AoA) method, and time difference of arrival method (TDOA).

Existing observed time difference (E-OTD) method [34] is one of the TDOA methods that becomes a de facto standard for U.S. Federal Communications Commission (FCC) ’s enhanced 911 (E911) service Phase II implementation[3] for GSM network. The accuracy of E-OTD is ranging from 50 meters to 125 meters because many factors including multipath fading and channel conditions can impact the accuracy of distance estimation. In addition, the E-OTD method has a slow response time, which is close to 5 seconds. In other cellular networks, there are also similar methods. For example, Wideband Code Division Multiple Access (WCDMA) network combines Observed Time Difference of Arrival (OTDOA) and idle period downlink (IPDL) for positioning.

However, triangle algorithm based positioning systems face some challenges. Compare with GPS, the accuracy of this kind of methods is low. In most cases, extra infrastructure has to be deployed and software on customer side need to be modified. For example, E-OTD requires synchronous network while GSM is not. To implement it in GSM network, operators have to build extra location measurement unit (LMU) that incurs extra cost.

2.2.3 Cell-ID aided positioning system

The third category of positioning system – Cell-ID based methods were recommended by 3rd Generation Partnership Project (3GPP) [38]. The idea is to simply retrieve and exploit cell sector information. This method does not require extra infrastructure and modification in software. But the weakness is also obvious. The accuracy of Cell-ID based method is strictly based on the size of cell sectors. Considering that the cell size is as large as 0.5-3 km in
urban areas and 3-20 kilometres in suburban areas, the accuracy of Cell-ID based approach is the worst among the categories of positioning systems.

On the other hand, the Cell-ID based approach is much more energy efficient than the other approaches. Thus, it is possible to adjust the trade-off between accuracy and power. In 2011, Paek et al. [38] proposed a Cell-ID Aided Positioning System (CAPS). CAPS uses a Cell-ID sequence matching technique to estimate position based on historical Cell-ID and GPS positioning sequences records. Paek et al. claimed that CAPS can reduce 90% total energy compared with GPS always-on case, and the accuracy is improved by 20% of the Android Cell-ID based positioning API.
Chapter 3

Signal collecting platform

This chapter will introduce the prototype of signal collecting platform. The first section explains the architecture and operation process. Then important technical details for each part will be introduced in the second section. The last section will discuss the extension possibility of the platform.

3.1 Architecture

![Figure 3.1: Architecture of signal collecting platform](image)

The architecture of the cellular network information collecting platform consists of three parts: back-end database whose role is to store information,
middle-ware server whose job is to process raw crowd-sourcing data and to provide Cell-ID historical data based service, and forefront client who responsibility is to act as information collector, and service providers/workers.

The distributed feature of the architecture provides good flexibility and security. First, although in the current prototype implementation, the database side uses a single mongoDB, it is possible to move it to the cloud for scalability. Second, the entire solution is based on crowd-sourcing, which means we are not able to control the content provider. Security of the whole system is of high priority. Thus, middle-ware plays an important role in preventing the database from being exposed to the external world. Furthermore, various secure mechanisms can be deployed in middle-ware to recognize those malicious data. For example, we can apply an asymmetric encryption scheme in the communication to filter malicious modified application and keep users’ privacy.

### 3.2 Database

In general, light-weight projects usually adopt relational database such as MySQL or Microsoft SQL Server as their storage solutions. However, in our prototype we choose the emerging NoSQL [32] database management system, i.e., mongoDB [8], based on the following advantages:

- Signal strength records in different places are independent, no complex relationship needs to be maintained in the data model;
- MongoDB is schema-free, the format of records can be dynamically modified without reprocessing the whole database;
- MongoDB is the best choice for distributed storage. It has better performance than SQL database for inserts, updates and simple queries[28];
- MongoDB offers a clear path for horizontal scalability.

Every record in MongoDB is a document. All documents are stored in collections. A collection is analogous to a table in relational databases, and is a group of related documents that have a set of shared common indexes.

Figure 3.2 shows common fields of record stored in this signal strength platform:

- `_id` : automatically generated primary key;
- `cid` : unique number used to identify each Base transceiver station (BTS);
CHAPTER 3. SIGNAL COLLECTING PLATFORM

17

CHAPTER 3. SIGNAL COLLECTING PLATFORM

Figure 3.2: Record format in database

- **lac**: location area code;
- **lon**: longitude value in World Geodetic System established in 1984 (WGS84) [13];
- **lat**: latitude value in WGS84 standard;
- **ss**: average signal strength;
- **cnt**: number of samples;
- **time**: updated timestamp of last sample.

Herein, the logical primary key is tuples lon, lat, lac, cid. Thus, if it is possible that the device can connect to different base stations at the same location, corresponding number of records will be generated in the database.

3.2.1 Data compression

Before updating records, two techniques have been applied to compress data and improve performance. First, GPS coordinates are truncated to 4 decimal places and converted to integers. In other words, the achieved granularity is approximately 5 meters. Second, for all data with the same logical primary key, the signal strength will be calculated by a smoothing function. The motivation behind this action is based on the experiments conducted by Ou et al. [26]. Figure 3.3 is the result of signal strength variation tests for two different phone models. Although we keep these phones static, the variation range of signal strength is still as high as 10 dBm due to reasons such as multi-path fading. In our prototype, we select the average function as the smoothing method.

After these two compression techniques, we merge those homogeneous data to save storage space. Up to now, the dataset that covers Otaniemi
area (Espoo, Finland) occupies only 200 MB. The size of the backup file that uses mongoDB’s built-in compression method is only 6 MB. This number demonstrates feasibility to store local signal strength data directly on smartphones.

### 3.2.2 Difference between models

Figure 3.4 reveals another fact that signal strength measured by different phone models are also different. This phenomenon is more obvious for
CHAPTER 3. SIGNAL COLLECTING PLATFORM

<table>
<thead>
<tr>
<th>Action</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upload signal information</td>
<td>[U#MAC#MODEL#LAC#CID#SS]</td>
</tr>
<tr>
<td>Request for position</td>
<td>[R#Nonce#MAC_1#MAC_2#..#MAC_n]</td>
</tr>
<tr>
<td>Response position to client</td>
<td>[P#rID#lat#lon#n#nord#maxValue]</td>
</tr>
</tbody>
</table>

Table 3.1: Communication Protocol

moving scenarios (Figure 3.4). Radio modules, types of antennas, handover algorithms and many other factors may impact the behaviours of smartphones. Due to this reason, the platform uses collections to classify data by smartphone’s models. Furthermore, this mechanism and the data compression technique mentioned above efficiently protect users’ privacy. Sensitive information such as MAC address and IP address is not recorded.

3.3 Middleware

Middleware is the most important component that acts as a bridge between data and user. It supports both TCP and UDP communication modes. In current implementation, the middleware is a Python server which runs at port 38921. The only service offered by the current implementation is a Cell-ID based positioning service. Table 3.1 defines the packet’s format for communication between middleware and client.

Recall that we mention the signal collecting process is anonymous. Thus, MAC address is optional while uploading. Nevertheless, in the prototype, MAC address is mandatory due to the requirement of “Handprint” system, which will be explained in the second part of the thesis.

3.4 Client-side

To demonstrate the operating process and feasibility of the platform, three different applications have been developed, i.e., collecting signal strength manually, collecting signal crowd-sourcing approach, and analysing data (refer to Figure 3.5).

3.4.1 Signal collecting

For bootstrapping the system, all crowd-sourcing based services face the same problem - lack of data. Hence, it is necessary to develop a manual signal-collecting application for bootstrapping the system. Figure 3.6 illustrates the
Figure 3.5: Three applications
CHAPTER 3. SIGNAL COLLECTING PLATFORM

Figure 3.6: App for manually collecting signal data

cellphoneModel
GT_I9505

Running Time
47 2014-03-27

GPS INFO
lat:60.18631092
lon:24.8200926

Debug Information
elisa(24405)
lac: 29120 cid: 286160 ss:-69

WiFi
No WiFi Connection
Figure 3.7: Helsinki Bus Timetable

app we developed to collect data by wardriving [18]. In the collecting process, this application enables GPS to obtain reference positions. After that, it uploads real-time cellular network information (e.g., signal strength, Cell ID) per one second. Furthermore, this application can automatically stop uploading when GPS becomes invalid, e.g., walking into indoor environment.

3.4.2 Helsinki Bus Timetable

The second application, i.e., Helsinki Bus Timetable (refer to Figure 3.7), demonstrates how crowd-sourcing approach works in a common LBS application. This application is designed for people to catch a bus. In this scenario, users who are familiar with the routes only want to know the accurate arriving time of the next bus. It usually takes several steps with traditional
map application: open the map, enter the start point, enter the destination and click “planning” button.

With Helsinki Bus Timetable app, users only need one or two clicks. The software will automatically locate the user, and display the timetable of the closest stop. The data chart (Figure 3.7) shows how it combines the Cell-ID aided positioning system with GPS. The crowd-sourcing platform and users both benefit from the application. When GPS is enabled, the data will contribute to the crowd-sourcing platform. Otherwise users can obtain an estimated location from the platform according to the cellular information. In this scenario, users do not need very high positioning accuracy. Cell-ID based positioning system brings many advantages compared with GPS: energy saving, indoor coverage and fast response.

3.4.3 Research app

To facilitate research and analysis, research apps are allowed to interact directly with back-end database (refer to Figure 3.5). This design isolates trusted and reliable researchers from potential malicious users. Even the middleware crashes, research apps still work. Figure 3.9 demonstrates one of the apps, i.e., “signal strength map”, which displays coloured points as the signal strength record on map. With this tool we find three interesting phenomena that will be introduced in the next chapter.

3.5 Extension

Mobile crowd-sensing is a new concept that was introduced by Raghu K. Ganti et al. [15] in 2011. Contemporary portable computing devices, including smartphones, embedded gaming gears, and in-vehicle sensing devices, are all equipped with multiple sensors. From the data collected from these devices, we can measure various individual and community phenomena, such as movement patterns (individual), modes of transportation (individual), pollution levels (community) and real time traffic patterns (community). Relevant solutions and methods are called crowd-sensing.

From the definition, we can see the idea of crowd-sourcing based signal strength platform is very similar to crowd-sensing. Hence our crowd-sourcing platform can be easily extended to cover crowd-sensing applications. For example, in order to collect the Wi-Fi fingerprints, we generate a new collection called “Wi-Fi” in the back-end database and slightly modify the communication protocol. After that we implement a Wi-Fi distribution map function that is similar to OpenSignal’s application. Signal strength collecting appli-
Figure 3.8: Data flow of Helsinki Bus Timetable
Figure 3.9: Signal Strength Map
cation and Helsinki Bus Timetable are analogous to crowd-sensing applications that fall into participatory sensing model [6] and opportunistic sensing model [22].
Chapter 4

Observations

After implementation of the prototype platform and the related applications, we deploy the signal collecting application, as mentioned in section 3.4.1, and collect data in Otaniemi area (Espoo, Finland) by walking and taking public transportation. Parameters of the devices under test are listed in Table 4.1.

The dataset covers most major streets and pedestrian routes of Otaniemi and Tapiola area (refer to Figure 4.1). Furthermore, parts of bus routes in Espoo is also included. Based on cellular information collected in this area, we find some interesting phenomena.

4.1 Density of base stations

Otaniemi is a suburban area in Espoo, Finland, where the main campus of famous Aalto University is located. However, in this small area whose diameter is approximately 2 kilometres, we collect more than 67 different Cell-IDs. Considering that all samples are collected by the same operator (Elisa), the total number of base stations deployed in Otaniemi will be even more if all the operators are taken into consideration. Generally speaking, we think the cell size is counted by kilometres, the density of cell towers here is much higher than expected.

<table>
<thead>
<tr>
<th>Phone</th>
<th>Model</th>
<th>Operator</th>
<th>Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung Galaxy S4</td>
<td>GT_I9505</td>
<td>Elisa</td>
<td>20831</td>
</tr>
<tr>
<td>Samsung Galaxy S2</td>
<td>GT_I9100</td>
<td>Elisa</td>
<td>14851</td>
</tr>
<tr>
<td>Samsung Galaxy Nexus</td>
<td>Galaxy_Nexus</td>
<td>Elisa</td>
<td>18746</td>
</tr>
</tbody>
</table>

Table 4.1: Statistics of collected data
Figure 4.1: Signal strength distribution in Otaniemi
Figure 4.2: Distribution of base stations
CHAPTER 4. OBSERVATIONS

4.2 Small cells

Further analysis of each Cell-ID’s coverage reveals more interesting facts. For a given Cell-ID, we can use the number of different records in our database to represent the size of “perceived coverage area”. Figure 4.2 depicts the size distribution of Cell-IDs’ “perceived coverage area”. More than half Otaniemi’s detected Cell-ID cover less than 200 points in the dataset, while large cells who have more than 1000 records also exist (refer to Figure 4.3).

We divide the cells into 5 types by their dimension (Table 4.2). Measurements in Otaniemi area show that most “perceived” cells are small macrocells or microcells. This observation reveals an important fact that has been neglected for a long time. The fact is the positioning error of Cell-ID based positioning schemes is reducing all the time with the explosive adoption of new telecommunication technology. It is time to revisit and re-evaluate the accuracy of Cell-ID based methods.

Table 4.2: Cell Types [35]

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Cell dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large macrocell</td>
<td>3-30</td>
</tr>
<tr>
<td>Small macrocell</td>
<td>1-3</td>
</tr>
<tr>
<td>Microcell</td>
<td>0.1-1</td>
</tr>
<tr>
<td>Picocell</td>
<td>0.01-0.1</td>
</tr>
<tr>
<td>Nanocell</td>
<td>0.001-0.01</td>
</tr>
</tbody>
</table>
CHAPTER 4. OBSERVATIONS

The situation in the future may be even better. Current networks already reach their capacity limits for both wireless access networks and backbone networks in highly populated metropolitan areas during peak times. On the other hand, SMART 2020 spurs environment friendly technologies in order to tackle the ever-increasing carbon emissions caused by information and communication technology (ICT) [36]. Cell-size reduction is the simplest and most effective way to increase system capacity and save energy. Jakob Hoydis et al. [17] proposed the concept of small-cell networks (SCNs) which is surrounded by the idea of cell-size reduction. Deploying dense self-organizing, low-cost, and low-power small cells to replace existing macrocells is a feasible direction for future telecommunication industry.

In addition, a large cell may be perceived as several small cells in a specific area due to many reasons. Figure 4.4 depicts an example for the phenomenon. In this case, buildings block signals in the middle of the road, hence cell
309543 can be divided into three sectors according to signal strength.

4.3 Intersection of multiple base stations

The third observation from Otaniemi’s dataset contradicts with network planning theory. It is well known that the planning of cellular network is based on hexagonal areas for efficiently reusing limited frequencies [29] [24] [30]. An optimal deployment of base stations in theory looks like a honeycomb (Figure 4.5). Base station locates in the center of each cell. Assume this is an ideal surface without obstructions and base stations have same technical parameters, a smartphone will always connect to the closest base station to get the strongest signal strength. In other words, given a position, the optimally connected Cell-ID is unique. Considering most of handover algorithms for mobile devices set a threshold value, we can conclude that a smartphone may randomly connect to different base stations only when it is close to the edge of a cell.

However, dataset in Otaniemi area presents another picture. We find out
that a smartphone may connect to more than two different towers in most area we measured. Figure 4.6 depicts the coverage of multiple base stations. In the figure, red point means in this position only one Cell-ID is detected. Blue point means there are two different Cell-IDs detected. Green represents positions where more than two different Cell-IDs have been detected.

4.4 Further observation of multiple coverage

In this section, we design several experiments to study possible factors that lead to the “multiple coverage phenomenon” mentioned above.

The first experiment aims to observe the behaviour of a static smartphone. We put a Samsung Galaxy S4 smartphone in one office in T-building (CS-building) of Otaniemi campus of Aalto University. Mobile data is enabled and no other application is running except system services. We implement a program to record the cell ID per one second for 30 minutes, as shown in Figure 4.7. The figure shows that four different Cell-IDs can be detected in this period and Galaxy S4 frequently changes its attachment among them. After that we conduct the same test on Samsung Galaxy S2 and observe
similar result. Their average handover time calculated from 10 hours’ data
is 24.21 seconds and 58.11 seconds, respectively.

A possible explanation is that this position is covered by several competitive base stations so that sensitive soft handover algorithm frequently changes channels to transmit background traffic. Soft handover is only valid under active state [21]. In order to verify this explanation, we repeat this experiment but disable the mobile data. As expected, the connected cell IDs become much more stable under idle state, as depicted in Figure 4.11.

After that, we design a second experiment to test the impact of another critical factor: the model of smartphone. In this experiment, two phones with different models are put at the same location, a test software records and compares their Cell-IDs per one second (Figure 4.8).

Figure 4.9, 4.10, and 4.11 are results of experiments tested under all three possible scenarios: active state v.s. idle state, active state v.s. active state, and idle state v.s. idle state. Each experiment lasts for more than 10 hours. We find out that Samsung Galaxy S4 and Samsung Galaxy S2 have completely different Cell-ID preferences in these experiments (Table 4.3).

These experiments are also evidence to prove how complex today’s cellular networks are. Possible reasons under these results include handover, physical properties of antenna, load balance in cellular networks, and different network types (GSM, CDMA etc.). All these factors make a smartphone have its own “personality”.

<table>
<thead>
<tr>
<th>Test</th>
<th>Total</th>
<th>Same</th>
<th>Pct.</th>
<th>Diff</th>
<th>Pct.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle Vs Idle</td>
<td>30617</td>
<td>10705</td>
<td>34.964%</td>
<td>19912</td>
<td>65.036%</td>
</tr>
<tr>
<td>Active Vs Active</td>
<td>61234</td>
<td>9670</td>
<td>15.79%</td>
<td>51564</td>
<td>84.208%</td>
</tr>
<tr>
<td>Active Vs Idle</td>
<td>126070</td>
<td>11000</td>
<td>8.725%</td>
<td>115070</td>
<td>91.275%</td>
</tr>
</tbody>
</table>

Table 4.3: Cell-ID preference
Figure 4.8: Compare Cell-ID
CHAPTER 4. OBSERVATIONS

Figure 4.9: Active vs Active

Figure 4.10: Active vs Idle

Figure 4.11: Idle vs Idle
Chapter 5

Handprint positioning system

In this chapter, we will introduce “Handprint”, an innovative Cell-ID aided positioning system that is inspired by the observations mentioned above. In the evaluation part, we will compare the Handprint system with existing API and the latest Cell-ID aided positioning system.

5.1 Background

There are two ways to exploit Cell-ID for positioning. The first one is sequence matching. Sequence matching methods store the Cell-ID sequence of traces. When users need positioning while they are moving, the method will compare the passing Cell-ID sequence with the pre-stored data. Researchers from Vietnam applied this method and effectively estimated location and velocity of buses [25]. Energy friendly positioning solution CAPS is also based on similar idea [27].

The second way addresses positioning problem for opposite scenario. Given a fixed position in a real world, the propagation channel is always unstable due to multipath effect etc. Fingerprint method measures and pre-handprint these multipath signal characteristics, and store them in a database. When a device requests positioning service, there must be a static measurement time period to collect enough information for pattern matching. Fingerprint based methods perform well in urban area and do not require additional equipment. Takenga et al. [33] proposed a system that considers the signal strength of nearby base stations as fingerprint. Simulation with Matlab shows decent accuracy in Hannover.

However, none of these solutions is compatible with both mobile and static scenarios. The core principle contradicts with each other: user must stay for a while to measure the multipath characteristics while this behaviour leads
to lack of Cell-ID sequence. For fingerprint system, extra scanning time to extract signals from different base stations results in waste in time and energy consumption, which is considered as a major advantage compared with GPS. Thus, the key point to improve performance of fingerprint methods is to make the sensing process instantaneous, and energy friendly.

5.2 Handprint

Borrow the idea from crowdsourcing, parallel computing and peer-to-peer network, “Handprint” system divides the fingerprint collecting process into many subtasks and allocates them to many devices. According to the observations mentioned in the previous chapter, we observe that phones may connect to different base stations at the same location. Even if they connect to the same base stations, they may have different signal strength. Considering the <cell-id, signal strength> pairs as the “Fingerprint” of a smartphone, the idea of “Handprint” is that several devices at the same location can share their fingerprint information to collectively determine the position (Figure 5.1). It makes sense because Cell-ID aided positioning systems are primarily designed for urban areas. In urban area there are usually many de-
VICES surround a specific user, e.g., when taking a bus, walking with friends, or waiting for a bus with other passengers. Sometimes even one user can own several types of devices, such as phones, tablets and wearable devices.

The operation process of Handprint system can be summarized as three steps: share fingerprint with nearby devices, send combined “Handprint” information to server, wait for response from the server. In this thesis, we integrate the prototype of “Handprint” system with the crowdsourcing platform mentioned above. All the computation is completed on the server side by middleware.

5.3 Implementation details

The implementation of the prototype faces two major challenges: the first one relates to how to share fingerprint information; and the second one is how to estimate position with these fingerprint. This section will first compare several mainstream short range technologies and then propose a position estimation algorithm.

5.3.1 How to share fingerprint information

Recently mainstream smartphones usually integrate multiple communication interfaces, including Bluetooth, IEEE 802.11 (Wi-Fi), and NFC.

Bluetooth is a wireless communication standard targeting at wireless personal area network (WPAN). It is used to replace cables and interconnects communication devices such as phones, keyboards, printers and mice. After the original idea was conceived in 1994 by Ericsson Mobile Communications, Bluetooth was soon widely adopted by industry and finally accepted by IEEE 802.15 as an official standard. Nowadays, almost every smartphone is equipped with Bluetooth.

Wi-Fi is another wireless communication standard that also works at ultra high frequency. However, Wi-Fi is originally designed to build up connections among computers. The range covered by Wi-Fi standard is called wireless local area network (WLAN). This concept is similar to WPAN. Consequently, products utilising Wi-Fi technology usually overlaps those Bluetooth devices. In general, Wi-Fi is also an indispensable component for smartphones.

The different features between Bluetooth and Wi-Fi are listed in Table 5.1. From the table, we can find that Bluetooth reduces power consumption at the expense of communication range and transmission rate. As a rule of thumb, Bluetooth is usually used to transfer files between smartphones while Wi-Fi is used to connect to the Internet because it can offer higher
TABLE 5.1: Comparison between Bluetooth and Wi-Fi

<table>
<thead>
<tr>
<th>Standard</th>
<th>Bluetooth</th>
<th>Wi-Fi</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE spec.</td>
<td>802.15.1</td>
<td>802.11a/b/g</td>
</tr>
<tr>
<td>Frequency band</td>
<td>2.4 GHz</td>
<td>2.4 GHz; 5 GHz</td>
</tr>
<tr>
<td>Max rate</td>
<td>1 Mb/s</td>
<td>54 Mb/s</td>
</tr>
<tr>
<td>Nominal range</td>
<td>10 m</td>
<td>100 m</td>
</tr>
<tr>
<td>Nominal TX power</td>
<td>0 - 10 dBm</td>
<td>15 - 20 dBm</td>
</tr>
<tr>
<td>Energy consumption</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

bandwidth than Bluetooth (as a matter of fact, even higher than cellular network).

For Handprint system, short communication range is not a weakness but a strength. In Handprint system, we assume all devices are located in a small area, share their cellular information to collaboratively estimate the position. Long distance among devices will reduce the value of the knowledge shared, decrease efficiency, and increase the complexity of positioning algorithms. Furthermore, there is no high-demand on transmission rate because the exchanged cellular information is as little as several bytes.

At a first glance, Bluetooth seems to be more suitable for Handprint due to its proper communication range and low power consumption. Nevertheless, we use Wi-Fi Direct [2] instead of Bluetooth due to the following weaknesses of Bluetooth. First, there is a long device discovering time that usually lasts for more than 10 seconds. This long discovering time will make Handprint lose response time advantage compared with GPS if it uses Bluetooth. Second, Bluetooth’s master/slave operational mode makes it difficult to instantly share informations among more than two devices. The last and fatal weakness of Bluetooth is that it defines a mandatory pairing process for authentication. It is cumbersome to ask verification code from strangers for a simple positioning purpose.

Thus, we choose Wi-Fi Direct as the communication technology to exchange information. Wi-Fi Direct was initially named Wi-Fi P2P, and is an emerging standard used to build connections between devices without access point. This standard is solely implemented by software, thus, it does not require extra hardware if a Wi-Fi module already exists in the smartphone. Compared with Bluetooth, the setup time of Wi-Fi Direct is less than 1 second. Nevertheless, it also has a similar feature with Bluetooth, i.e., it has a pairing process before exchanging information.

To address the pairing problem, we utilise an indirect way to share cellular information. At the beginning, every smartphone equipped with Handprint
system must maintain a connection between itself and a centralised server. After the connection is set up, the smartphone will periodically broadcast a Wi-Fi Direct service beacon. If a device needs to estimate its position, it can discover nearby devices by extracting MAC address from the beacons it received. Then it will send a request that contains the MAC address list to the centralised server. The centralised server is responsible for calculating the estimated position. If there is no cellular information for a device, it will send a notification to the corresponding device. Figure 5.2 describes a flow chart of the cellular information sharing process. This design avoids pairing process while the event-driven mode keeps the radio on idle state when there is no positioning request.

Besides communication between smartphones, there are many other types of wireless communication networks such as wireless sensor network. The possibility to extend Handprint system to other networks will be discussed in Chapter 6.

Figure 5.2: Flow chart of cellular information sharing process
5.3.2 How to estimate position

Handprint system adopts an iterative algorithm to estimate the position. Due to factors like multipath propagation, signal strength of a moving smartphone is not as constant as that of a static one, occasionally the variation is as large as 10 dBm. Thus, we introduce a tolerance parameter, i.e., $\epsilon$, to reflect this fact. When the server receives a positioning request, it will first look for potential positions for each device in range $[ss-\epsilon, ss+\epsilon]$. After that, it will check the number of the points belonging to the intersection of all sets. If the number of points is larger than a given threshold $C$, the algorithm will decrease $\epsilon$ and calculate again. When the loop ends, the estimated position will be returned by calculating weighted average value of all intersection points.

\begin{algorithm}
\textbf{Data:} Cellular information list $[model_i, cid_i, ss_i]$
\textbf{Result:} Estimated position \\
Set threshold $\epsilon$ and $C$;
\textbf{while} size of $Set_{int} < C$ \textbf{do}
\begin{itemize}
  \item $Set_i = \{\text{position which satisfy } cid = cid_i \text{ and } ss_i - \epsilon < ss < ss_i + \epsilon\}$
  \item $Set_{int} = \text{intersection of all } Set_i$
  \item $\epsilon = \epsilon/2$
\end{itemize}
\textbf{end}
\text{Calculate weighted average value of coordinates in $Set_{int}$;}
\end{algorithm}

Furthermore, if $Set_i$ is an empty set in the computing process, Handprint will drop it. This technique is designed to increase robustness. The logic behind this action is that if there is no knowledge about this device (no one has uploaded it before), we ignore it. Thus, if there is only one valid device, Handprint will fall back to a generic Cell-ID and signal strength based solution.

5.4 Performance evaluation

In this section we evaluate our Handprint system, in comparison with two existing Cell-ID aided positioning systems.

The first reference positioning system is Google Geolocation API. Google Geolocation API is integrated into Android smartphones. It is a fused method that combines fingerprint of Wi-Fi and Bluetooth devices, location of cell towers, and positioning result from GPS. With this API, Google Map can estimate location for users even if GPS is disabled. Undoubtedly, GPS
The second reference system is CAPS [27], a Cell-ID aided positioning system that leverages near-continuous mobility and the historical positioning information of a user to achieve better accuracy than Cell-ID based approach. Experiments in Korea and U.S. demonstrated that it is energy efficient under acceptable accuracy.

5.4.1 Experiment scenario

In the experiment, we assume a user is taking a bus from Westendin asema to Leppavaara, Espoo, as shown in Figure 5.3. In this scenario, the Handprint positioning system is used for navigation, while GPS is used as a ground truth to calculate accuracy of the Handprint system.
This area can be considered as a suburban area that has median population density. Experiments on this route can be considered as a lower bound for most urban areas in Finland, because there are more base stations in Helsinki urban area.

All data collected with Handprint are collected with two smart phones, which is a minimal requirement of this system. Otherwise the Handprint system will be fallen back to a generic Cell-ID based scheme, as mentioned before. In principle, the more devices participating in the Handprint system, the more accurate positioning can be achieved. Thus, we can consider the following experimental results are lower bound.

### 5.4.2 Compare with CAPS

We use one Samsung Galaxy S4 and one Samsung Galaxy S2 as the devices under test. In the 20 minutes journey, requests are sent per five seconds. Test results are listed in Table 5.2. From the table, we can see that 96.8% of requests receive valid responses and the positioning error of more than 60% of samples is less than 100 meters on both devices.

CAPS is not open source, thus, we are not able to test it directly on the same route. However, we can compare the result that was published in their paper [27]. Figure 5.4 depicts the CDF graphs of Handprint and CAPS. The data of CAPS is collected at Los Altos, U.S. This route has similar length and is tested under similar driving speed with the route we choose in Finland.

Experimental data reveal that although the median error of Handprint system is close to that of CAPS, the 80th percentile error of Handprint decreases by 26.3% compared with CAPS. The red line in the right figure represents the CDF graph of positioning error of pure Cell-ID based approach.

### 5.4.3 Compare with Google Geolocation API

Google Geolocation API is the existing most popular assisted positioning system that has been widely applied to Android smartphones. After receiving a positioning request, Google Geolocation API will return a location and
accuracy radius based on information from cell towers and Wi-Fi nodes that the mobile client can detect.

We repeat the same experiment with Google Geolocation API on the route mentioned in Figure 5.3. The average error is 239 meters while the median error is 174.42 meters. Both of these two numbers are significantly higher than that of Handprint system. To analyse the variations of positioning error, we plot the error per one second in Figure 5.5. Through this figure, we can find that in some locations the error jumps to more than 1000 meters. This can be explained by large cell towers, as shown in Figure 4.4. Furthermore, we also observe that the frequency of Google Geolocation API to update positions is much lower than the request frequency. This is because in a short period of time (e.g., tens of seconds), the cellular information barely changes. Google Geolocation API can not update the position because it does not obtain any new information.

### 5.5 Converging speed

Another important factor will needs to be taken into consideration is the converging speed. Namely, how many data we need to collect by crowd-sourcing to make Handprint work efficiently? To analyse this question, we choose another trace that starts at Sello and ends at Ikea in Espoo, Finland, as shown in Figure 5.6. This route is a motorway. Supposedly, it is different from the route we tested in the previous sections, which is a regular arterial road.
Figure 5.5: Error of Google Geolocation API

Figure 5.6: Chosen route of Bus 270
One Samsung Galaxy S4 and one Samsung Galaxy S3 are used to collect cellular information along this road. These two phones use network services from two different operators. In this experiment, we collect 10 trace files for each smartphone. Each time a pair of new record files are uploaded, we simulate taking a bus with two phones, and sending positioning request per one second.

Figure 5.7 illustrates the distribution of response types. “Invalid” response means there are no corresponding data so Handprint can not estimate the location. “One phone” means the estimated position is calculated with information from one device. “Two phones” means the result is calculated from the intersection of two phones’ possible position sets. According to these definitions, the percentages occupied by “Two phones” represent the state of Handprint. We can see from the figure that after ten traces are uploaded for each device, Handprint is available for 90% of requests, and 60% of response is calculated by Handprint algorithm. As a summary, 10 traces are enough to make Handprint system work.
5.6 Power consumption

The power consumption of Handprint system is divided into two parts: the processing of exchange information, and the communication between the client and the centralised server.

Although for each request, Handprint only sends several bytes to the server. The power consumption of this process is not negligible because for 3G smartphones, 60% of energy is wasted by tail energy [4]. Thus, the tail energy incurred by sending the requests is an overhead for Handprint system. This overhead exists if the network traffic of Handprint occurs when the radio is idle. Nevertheless, considering that positioning services offered by Handprint are used for online location based services in most cases, this overhead is amortised by other applications. Furthermore, according to the event-driven mechanism described in Figure 5.2, mobile devices do not need to upload its cellular information if there is no other device close by. Finally, the analysis in Chapter 3 also illustrates that it is feasible to develop an off-line version of Handprint to further decrease tail energy.

The process of exchanging information is also energy efficient. First, most of short range communication protocols such as Bluetooth takes into consideration energy consumption from the beginning. Second, users usually enable at least one of them in regular daily usage scenario. Current implementation of Handprint only detects the MAC address of neighbouring devices that may be extracted from regular Wi-Fi scanning or other communication techniques.

In brief, the energy consumption of Handprint is shared by regular applications and usage scenarios. Handprint is an energy efficient solution compared with GPS.
Chapter 6

Discussion and future work

In this chapter, we will discuss some future work of the crowdsourcing platform and Handprint positioning system.

6.1 Knowledge Share Protocol

Review the whole implementation of Handprint system, we can find that the key contribution is to attract attentions on short range ad hoc networks consisting of smart phones. Conventional location based services and positioning methods have not taken into consideration making collaborative decision with other smartphones.

The core idea of Handprint, let us name it “P2P network in reality”, is similar to the concept of ubiquitous computing. One of the most potential future work of Handprint is to extend it to ubiquitous network. In other words, we can design a knowledge share protocol to share geolocation knowledge among smartphones, access points, sensor networks, NFC devices and all other types of wireless devices. For example, we can manually label geolocation coordinates for a sensor network, every nearby user can then receive the position information that is encapsulated by Knowledge Share Protocol, as shown in Figure 6.1.

It is also possible to utilise Knowledge Share Protocol to combine Handprint with other positioning system. As shown in Figure 6.2, for several users near each other, only one user needs to enable GPS. In this situation other users can still obtain accurate position by Knowledge Share Protocol. For large city such as New York or Shanghai, similar scenarios appear every day.

Knowledge Share Protocol has many advantages. From user’s perspective, positioning by Knowledge Share Protocol is available for both indoor and outdoor scenario. The decreased acquisition time of GPS brings energy
CHAPTER 6. DISCUSSION AND FUTURE WORK

Figure 6.1: Extend to sensor network

Sensor network, broadcast (lat, lon)

Figure 6.2: Share information from GPS
saving and fast response time. From social perspective, it significantly saves energy due to the low power consumption feature of short range communications. For instance, in Figure 6.2, we can use short range communication to replace GPS for the rest of users except the provider.

6.2 Future improvements on the prototype

Currently the prototype of crowd-sourcing platform and Handprint system are still in lab environment. Future work can be done from the following aspects, to make it more realistic.

First, there needs a real popular application to help collect data by crowdsourcing. To support large scale storage, modification on the architecture of crowdsourcing platform shall be executed to support distributed storage.

Second, to verify the interesting phenomena we observed and to improve the performance of Handprint system, more experiments on other locations shall be done, e.g., city centres. Unfortunately, to achieve this goal we need to collect enough data first. Nevertheless, the platform supports wardriving to collect data. Thus, it is easy for other researchers to manually repeat these experiments. The source code of the whole project will be open-sourced and be accessible from Internet.

Third, in this thesis, all experiments of Handprint system are completed by two smartphones. It is be interesting to see whether the accuracy will be increased by using more smartphones. In this case, we need to think about a new topic: robustness. At the early stage when there is not enough data for a model in database, the positioning result may be as worse as several kilometers. Such a device may pollute other devices if it broadcasts inaccurate result. Thus, a mathematical model must be built to evaluate the reliability of data. For example, we can use numbers to represent trustiness. The trustiness of location obtained by GPS can be set as 1, for Handprint location estimated from reliable model is 2, and for model lack of data is 3.

Finally, details of Handprint prototype can also be improved. Current version uses TCP to maintain the connection. We can change it to UDP In the future. Beyond that, more authentication mechanisms can increase the security and prevent hackers to fake geolocation messages.
Chapter 7

Conclusions

In this thesis, we design and implement a cellular information collecting platform for increasing demands from various application and researches. As a successful application of this platform, we invent a new energy efficient solution for positioning.

The contributions of this thesis are summarised as follows:

- We have designed and implemented a crowd-sourcing platform that is used to collect cellular information such as Cell-ID and signal strength. This platform has flexible architecture and good extensibility.
- Based on data collected in Espoo, area, we found out there interesting phenomena of existing cellular networks. The density of cell towers and number of “Small cells” is increasing. Furthermore, most areas in Espoo are covered by more than one cell tower.
- Inspired by the phenomena as mentioned above, we design an innovative energy efficient Cell-ID aided positioning system called “Handprint”. It decreases 20% error comparing with latest research in academy and existing product from industry.
Bibliography


[12] Consumption, L. P. Location-based services on mobile phones: Minimizing power consumption.


