UiO : Department of Informatics
University of Oslo

Using Bluetooth beacons to pay for metro tickets

A case study in user control of ubiquitous computing

Torkjel Skjetne

Master's Thesis Spring 2016

Abstract

This master's thesis investigates the users' understanding, use and perception of control in ubiquitous and context-aware systems. An Android prototype was developed and tested with users. The prototype scanned for Bluetooth beacons and automatically purchased a metro ticket when a beacon was within close range. Automation of tasks aided by beacons was found to be fairly liberating, however, the participants felt unease when money was involved and felt they needed a higher degree of control. The application prompting the user to confirm the ticket purchase was preferred over an option to cancel an already purchased ticket. This thesis concludes that the transition to an environment of ubiquitous computing systems needs to be exactly that - a transition. Until the users can get accustomed to the technology and automation, a fair bit of user control is still desired, at least when automating monetary operations.

Key words: ubiquitous computing, context-awareness, user control, perceived control, mental models, in-situ testing, field testing, lab testing

Acknowledgements

I would like to thank my supervisor Jo Herstad for guiding me through all stages of this thesis, encouraging me and offering very constructive feedback. I would also like to thank John Olav Olsen at PayEx who helped me choose a thesis topic I thoroughly enjoyed, as well as giving me easy access to the beacons I used in my research. Additionally, a huge thanks goes out to the participants of my research who took time out of their days to provide me with data. Lastly, I would like to thank my friends and family who wanted to discuss my thesis with me and helped keep my spirits up.

Contents

List of Figures

Chapter 1

Introduction

Electronic devices and computers are becoming more prevalent throughout society and our everyday lives, filling roles in trivial tasks such as doing the laundry, driving a car, watching television or even sleeping. When we are surrounded by large and small devices interacting with each other, issues regarding trust, security and user control arise. Will the device really do what it's supposed to, will we feel like we're still in control, and will our personal information be kept hidden from unauthorized entities?

1.1 Background

Payment of goods and services is rapidly evolving from the use of cash, credit and debit cards using magnetic stripes or chip and Personal Identification Number (PIN) codes, through systems that use Radio Frequency Identification (RFID) or Near field communication (NFC) chips integrated into debit or Subscriber Identity Module (SIM) cards, and - more recently - to the use of Bluetooth Low Energy (BLE) beacons. When using NFC chips, the user holds his device up close to the payment terminal to initiate the transaction. Systems using technology such as BLE beacons, on the other hand, have a fairly large range. The trade-off, however, is that the short-range technologies such as NFC have a more precise location estimation. With the larger range and less precise location estimation of BLE, devices might behave in seemingly strange ways, executing automated actions not intended by the user. But these systems also help introduce a new era of interaction. As the user moves through the environment, systems react to environmental changes. The system has information on for instance where the user is, what time of day it is, what the user does, and so on. The systems sense and react to change of context - they are context-aware. They are said to "know" the context of the user. Whether the use of automated context-aware systems is understandable to the users, whether or not the users have a mental model that can explain the system's behavior, and whether or not they trust and feel comfortable with an automated payment system are questions that need answering when we move into the field of ubiquitous computing, context-aware systems and automated payment systems.

The use of Bluetooth beacons for location sensitive systems has been suggested and discussed on several occasions, e.g GSMA (2014), Salas (2014), and Faragher and Harle (2014). The many different use cases that accompany beacons have been exemplified, such as tracking industry supply chains, mapping customer movement and interest within museums or parks (SmartCardAlliance, 2014), or guiding customers through buildings (Vlugt, 2013). Others discuss how indoor location tracking can be accomplished from a mathematical and technological point of view (Bahl & Padmanabhan, 2000), (Roos, Myllymäki, Tirri, Misikangas, & Sievänen, 2002), (Subhan, Hasbullah, Rozyyev, & Bakhsh, 2011). However, research into whether or not the users will be able to put their trust in and perceive adequate control over automatized context-aware systems are of a more limited extent. The research that does focus on user control in ubiquitous computing has mostly been of a quantitative (e.g. (Spiekermann, 2005)) or mixed (triangulation) (e.g. (Barkhuus & Dey, 2003)) nature. Conducting qualitative research is important in the field of ubiquitous computing as it contains fairly new technologies and hence requires a close examination of the users' attitudes and experiences. As Kuniavsky explains it:

> *The more you watch people, ask them questions, and analyze their behavior, the more attuned you can make the experiences you design to their ways of experiencing the world. (Kuniavsky, 2010, p. 202)*

Therefore, this thesis focuses on the user's point of view, detailing, in a qualitative way, the user experience of an automated context-aware system and discussing the issues that arise.

1.2 Motivation

I have for as long as I can remember looked forward to a futuristic society where the newest technology is integrated with our lives. A society where everyday tasks and chores are as automated as they can be. Where I don't have to fumble around in my pocket to find my grocery list, or maybe where I don't have to go to the grocery store at all. Instead, my refrigerator tells the store when I'm out of milk, and the store comes to me. We are moving towards such a world, although not without hitches.

Not long before I started writing this thesis I managed to dent my car, initiated by an erroneous action from a context-aware system. I was leaving a dinner party at my parents' house, and when I turned on my car I suddenly heard a voice over the car speakers. I looked at the car's display and saw that a phone call was being received over Bluetooth. I had no idea what was going on, so I quickly hung up the phone using a button on the car dashboard. A few seconds later I heard the same voice again: "Hello? What happened?". It dawned on me that my father was sitting upstairs in his apartment, talking on the phone with his friend, and because he had set his phone to automatically connect to the car when within range, I received the call when I turned on the car. In a state of semi-panic I tried to back out of the driveway as fast as possible not to ruin their conversation. As a result, I bumped into a rock, ending up with the first ever dent in my car.

This is an example of how automation can lead to a loss of control and unwanted system behavior. After listening to a lecture by and later having a conversation with John Olav Olsen from PayEx[1] in Norway, I quickly came to the realization that Bluetooth beacons and automation of payment would be the perfect thesis subject for me. It's a field that is fairly unknown and untouched by the general public, and where I thought that a loss of control could be of great importance. And thus, I started working on this thesis, examining how users indeed understand and feel about the automated world that I so much look forward to. Are we ready for an automated world?

1.3 Research questions

The research questions of this thesis are related to the use and understanding of ubiquitous systems, specifically context-aware automatic payment systems. A system might not be understood or used to its full potential or as it was intended by the creators, because the users have a different mental model than what the creators envisioned. Hence, their mental models are not in line with the conceptual model the system was built upon. This might affect their perception and understanding of the system, hence their desire to use it. Additionally, will the users feel confident in applications making payments for them, or will the loss of control be too intimidating? The three research questions are as follows:

1. **How do the users' mental models of Bluetooth affect the use of beacon technology for automated context-aware payment?**

   Mental models affect how people handle situations and use systems. Many people have had some experience with classic Bluetooth, which although is somewhat different to the Bluetooth low energy (BLE) technology that beacons use. When enabling Bluetooth on a phone, the user does not choose to only enable BLE or classic Bluetooth. So in order to enable BLE, the user normally enables both Bluetooth technologies in one, not BLE specifically. Enabling Bluetooth is a prerequisite for location estimation using Bluetooth beacons, and so the users' mental models of Bluetooth might affect their desire to start using the system.

[1] www.payex.no

2. **What, if any, are the issues users face with regards to perceived control when using automated context-aware payment?**

   If the users don't feel like they are in control, they can feel inclined to abandon an application. Being able to trust that the application does what it's supposed to do, without any surprises, especially when money is involved, is an important factor for users in deciding whether to use a system.

3. **In what way does testing a ubiquitous computing system in the field result in different findings than testing in a lab?**

   Systems can be tested in a lab instead of in their intended use context, but at the cost of realism. The participants might not feel as immersed, won't be affected by other people or the general surroundings, or experience how stress that a realistic situation can induce might affect their use of the system. There are, however, difficulties with conducting user tests in context as well. Observing and recording exactly what the user is doing can be difficult, and the participants can feel uncomfortable if being asked to think aloud in public. Does testing ubiquitous systems in context really make a difference?

In this thesis I attempt to answer these research questions by discussing theory combined with findings from my chosen methods: conducting user tests of an Android prototype both in a lab and on the metro accompanied by interviews, and later conducting a focus group with a new set of participants. The prototype, which was developed by me, automatically purchases tickets for public transportation when the user gets close to a Bluetooth low energy beacon.

1.4 Thesis structure

Chapter 1 - Introduction The first chapter introduces the research area of the thesis, and explains the thesis' motivation as well as the research questions it discusses.

Chapter 2 - Case The second chapter introduces the case that the thesis is based upon, explaining the "Mobile applikasjoner underveis" research project that inspired the thesis, and describes Ruter - the agency for public transport in Oslo and Akershus.

Chapter 3 - Theory The third chapter explains the theory that the thesis and its discussion is based on. User experience, perception of control, ubiquitous computing and the act of connecting the physical world to the technological world are important areas of theory in this thesis.

Chapter 4 - Methods The fourth chapter describes the methods used in the thesis to answer the research questions. A fairly basic Android application functionality was developed as a prototype. This prototype was then tested in order to get the users' feedback and reactions. After these user tests the participants were interviewed semi-structurally. Lastly, a focus group with a new set of participants was conducted based on the research questions and topics encountered during user testing and interviewing, discussing the participants' experience and attitudes towards automation and Bluetooth.

Chapter 5 - Prototype The fifth chapter describes the development process of the prototype, mentions which beacons were used, as well as the functionality and minimalist design of the Android prototype.

Chapter 6 - Findings The sixth chapter details the findings from applying the methods, i.e. the user test, interviews and focus group, to the case using the prototype.

Chapter 7 - Discussion The seventh chapter discusses the research questions based on the theory and the findings.

Chapter 8 - Conclusion The eight chapter summarizes and concludes the thesis' research questions and discussion. Lastly, I talk about how others can build upon and continue my work from this thesis.

Chapter 2

Case

This chapter introduces the case that the thesis is based upon, explaining the "Mobile applikasjoner underveis" research project that inspired the thesis, and describes Ruter - the agency for public transport in Oslo and Akershus.

2.1 Mobile applikasjoner underveis

Institute of Transport Economics (TØI) is leading a research project called "Mobile applikasjoner underveis" (English: Mobile Applications Underway) which focuses on how mobile applications can contribute to making public transportation more attractive. The goal of their research project is to allow public transportation to compete with the passenger car. Areas of use for new applications are explored and tested, and information about the travelers' interest in and willingness to use mobile technologies is gathered. The project is a collaborative effort between Institute of Transport Economics (TØI), Viktoria Swedish ICT, Institute of Informatics at the University of Oslo, and Norges Statsbaner AS (NSB). Additionally, the project is being supported by Transnova, Statens Vegvesen, Flytoget and Ruter (Julsrud, Denstadli, Herstad, Hjalmarsson, & Li, 2014).

The uprising of fast and pervasive Internet access through the mobile network and WiFi allow mobile technologies to be used on the go, offering the use of a wide variety of applications that can not necessarily be feasibly used while driving a car. Julsrud et al. (2014) conducted a study of travelers in Oslo and Trondheim, showing that 98 % of travelers brought either their smartphone, tablet or laptop, while more than 80 % of public transport travelers brought their smartphone, showing that platforms for such applications are already in heavy use. One of the main focuses of the "Mobile applikasjoner underveis" project is user centered design, and bringing the users in at an early stage of development in order to research and develop what the users truly want and need.

My research on using Bluetooth beacons as an automated payment option for public transportation can help take some workload off the travelers. This allows them to do other things, such as reading those emails from work you were supposed to read yesterday or chat with friends, while not having to switch contexts just to pay for or validate your ticket. It's an application that enables further use of other applications and allows the users to focus their time on the applications they want. Julsrud et. al established four different groups of users of technology during travel:

1. Available underway
   The biggest group, consisting of almost 1/3 of all the users. The users in this group have access to smartphones and are experienced with the use of mobile Internet connections. They use their smartphones mostly for phone calls, text messages and some web surfing. They often drive cars on daily travels.

2. Work and news
   An active and technologically equipped group consisting of about 1/4 of all the users. During travel they mostly use technology for more functional activities such as writing or reading emails, text messages or reading news articles, while using a limited amount of social media and web surfing. This group consists mostly of men who often drive cars.

3. Active and social
   Just as the previous group, this group consists of about 1/4 of all

the users. These users, however, are younger and often use public transport. During travel they use their mobile phones or tablets to go on with social media, listen to music or for other entertainment purposes.

4. Technologically independent
   The last group consists of about 1/5 of all the users. They possess phones that often are without Internet access or they are without mobile technologies at all. When they first do use their phones or tablets it is to make phone calls or write or read text messages. While traveling, these users mainly read books or are busy with their own thoughts. This group mainly consists of elderly men.

An automatic ticket purchase application, if the users can put their trust in it, will allow the users in the three first user groups to focus on other tasks than that of purchasing a ticket. This will allow them to travel more seamlessly, eliminating any context switch that comes with manually purchasing a ticket. Hence, it can open up for uninterrupted use of social media, reading and writing of emails, or any other attention demanding activity. The people in the first two groups often travel by car, but enabling more seamless travel and ease of use can help make public transportation more attractive, hopefully reducing their daily use of cars.

2.2 Ruter AS

Ruter AS is the agency for public transport in the counties of Oslo and Akershus, where 60 % of the shares are owned by Oslo and 40 % owned by Akershus. Ruter is responsible for planning, developing, marketing and otherwise administering the public transport network in Oslo and Akershus. Ruter does not operate the public transport themselves, however, but they have contracts and agreements with different operators for buses, trams, metro and ferries. The local and regional trains are administrated by Norwegian State Railways (NSB). However, NSB and Ruter are cooperating in order to offer more seamless travel options between trains and the other public transport systems. As a result, for public transport other than trains Ruter is in charge of selling tickets, making public transport information available, and for customer support (Ruter, 2014). Ruter has released a mobile application called RuterBillett for purchasing single or season tickets, a mobile application called RuterReise which shows the public transport schedules and provides real time updates, and the website www.ruter.no with similar information.

In November 2015 Ruter sent out an open invitation asking for ideas on and offers to implement new payment options for their public transport tickets. They stated that "easily accessible and simple payment solutions constitute an important factor in reducing the barriers of using public transport" (Ruter, 2015).

2.3 Ruter's ticket system

A Ruter ticket can be purchased in the form of a single use disposable paper travelcard, an electronic travelcard for use more than once, or on a smartphone application. The electronic travelcards can be purchased online and mailed to the user, or they can be purchased in select stores or kiosks around Oslo and Akershus. The disposable travelcards can be purchased at ticket machines on many stations, or in the same stores and kiosks as the electronic travelcards[1]. If the user already possesses an electronic travelcard, the ticket machines can be used to purchase a ticket on that card. Single tickets can also be purchased from the bus driver or ticketing personnel on ferries, however, these are more expensive than if bought at a ticket machine or in a kiosk. On the mobile application RuterBillett any type of ticket can be purchased. The user is herself responsible for being able to present a valid ticket when the ticket inspectors arrive[2]. Different parts of the area Ruter is administering is split into different "zones". The ticket prices for travel within a single zone are the same for each zone, however, if you want to travel between zones you need a supplementary ticket. In 2011 Ruter introduced

[1]Sales outlets: www.ruter.no/en/buying-tickets/sales-outlets
[2]www.ruter.no/en/about-ruter/terms-and-conditions/transport-regulations

a new zone map, reducing the number of zones from 77 to 8, in order to move towards a shared system that would be easier to use in both Oslo and Akershus (Ruter, 2013). Single tickets last for one hour, and an additional half hour per zone you're crossing. Other ticket types are the 24-hour ticket, 7-day ticket, 30-day ticket, 365-day ticket, and special tickets for large groups or school classes.

In 2013 the National Institute for Consumer Research (SIFO) conducted research detailing the use of, experiences with and feelings towards the use of what they call "e-tickets" throughout Norway. They defined e-tickets as electronic contactless tickets based on NFC or QR, such as Ruter's travelcards. This definition of e-tickets does not include tickets purchased through mobile applications.Their research shows that, in 2013, 35 % of the respondents used e-tickets, and students were shown to be the most frequent users being represented with 53 %. In Oslo, a big part of Ruter's area of service, 56 % of the respondents used e-tickets. Lastly, it was shown that 54 % of smartphone users had one or more ticket apps on their phones. In the age group 18-29 years 68 % of the respondents had one or more ticket apps, compared to 33 % in the age group above 60 years (Slettemeås, 2014).

Chapter 3

Theory

This chapter outlines the theory that the thesis and discussion is based on. User experience, ubiquitous computing and the act of connecting the physical world to the technological world are the main areas of concern.

3.1 Human-computer interaction

The field of Human-Computer Interaction (HCI) studies how humans interact with computers, focusing on improving usability, usefulness and general user satisfaction. Usefulness is important, as a system can be easy to use, easy to learn, and have low rates of error, but it also needs to be useful to the users - it needs to serve them a purpose. HCI is an interdisciplinary field encompassing among others computer science, ergonomics, engineering, psychology, sociology and design (Shneiderman & Plaisant, 2010, p. 22).

During the 19th century, machines, most notably looms and sowing machines, were programmable and interactable through the use of punch cards - cards with holes in different locations, telling the machines how to operate. With the advent of the electronic computer, user interaction moved to the Command-Line Interface (CLI), where the computer presented the user with text on the screen, and in turn the user gave the computer commands in the form of text. In 1963, however, Ivan Sutherland developed a program called Sketchpad which allowed the user to draw and manipulate objects on the screen with the use of a light pen. Sutherland's Sketchpad is often viewed as the starting point of what was to become the Graphical User Interface (GUI) and the next stage of user interaction (B. A. Myers, 1998). The mouse, an integral input device for the GUI, was first developed in 1965, but was not commercially available before the early 1980's when released alongside systems such as the Apple Macintosh. During the early 1990's the mobile computing scene started to evolve quickly, and Personal Digital Assistants (PDAs) were in the hands of the consumer. As hardware was improved and new designs were made, increasingly functional mobile devices were accessible to the consumer, allowing them to carry small computers, phones and music players with them wherever they went. The GUI needed to be displayed on a small screen with no input device other than a stylus or a finger on a touch screen. As the way humans interact with and understand computers has evolved over the years, so has the goals and methods of HCI.

Harrison, Tatar, and Sengers (2007) argue that there are three different paradigms of HCI. The first paradigm they call a merge of *human factors and engineering*, that emerged from the need to reduce errors in airplane cockpits. It's concerned with solving problems in design, and one design is a solution to problems from previous designs found during use.

The second paradigm, which they call the cognitive revolution, focus on effective information processing. Here, design is mostly about efficiently and accurately communicating information between the machine and the users.

The third paradigm they call situated perspectives, focusing on how humans understand the interdependence of systems and how they can exist in different situations or contexts. This third paradigm encompasses the area of mobile, context-aware and ubiquitous computing that our society is now moving into, where it is important to understand that systems do not exist in isolation.

In the following sections I will further detail notion of this third paradigm, starting with an elaboration of ubiquitous computing.

3.1.1 Ubiquitous computing

In 1988 Mark Weiser talked about the next stage for computing, later referred to as the third wave of computing, in which we surround ourselves with computers throughout our environments where they aid us through everyday chores and activities (Weiser, 1991). The waves of computing are closely related, but not entirely equal, to the paradigms of HCI. Moving from the first wave of computing where large mainframe computers were used by many people, through the second wave where personal computers were used by a single or very few users, we now march forward to a third paradigm where the technology spread throughout the environment is adaptive, can interoperate, and can be used by many (Weiser & Brown, 1996). That is, we have moved from the one-to-many era through the one-to-one era and are now starting to scratch the surface of a new era of computing, many-to-many with interoperable and adaptive computers scattered throughout our environments. This next stage, called ubiquitous computing (or ubicomp for short), can be seen as differentiating itself from the traditional mobile computing through two main characteristics: physical integration and spontaneous interoperation (Kindberg & Fox, 2002).

A completely new discipline or field of study like ubiquitous computing offers new capabilities and opportunities, however it also brings its own problems, areas of concern and terminology.

Physical integration

In ubiquitous computing the devices and systems are related and connected to the physical world in some way. Ubiquitous systems are thought to often operate within environments already known from our everyday lives such as a living room, a kitchen, a car, a university campus, a classroom et cetera. These environments have to be equipped with different kinds of sensors and actuators to suit the needs of the systems that are expected to provide services in the environment. More detailed examples of how this can be done can be found in section 3.2, where I elaborate on the use of QR codes, RFID and NFC chips, and Bluetooth beacons.

Spontaneous interoperation

In the ubiquitous wave, the environments consist of both fixed components that are more or less always present, and transient or non-persistent components that spontaneously join and leave the environment. The component moving into or leaving the environment, as well as the components already in the environment, has to adapt and react accordingly. For instance, if a device in an environment routinely transmits data to other devices in the environment, the transmitting device has to stop attempting to transmit data to a device that leaves the environment. Additionally, the device leaving has to stop listening to data from the transmitter. And if a person enters a meeting room his phone can automatically switch to silent or vibrate mode, and the meeting room can update the participants list and tell the receptionist to bring him a cup of coffee. Or maybe someone wants to present something on the projector and give the meeting pariticpants notes for each slide for them to read on their phones. If an employee leaves the meeting room he/she does not want to receive the notes any longer. In the world of ubiquitous computing, this is ideally achieved without the need to install new software. Kindberg and Fox (2002) call this the Volatility Principle:

> *You should design ubicomp systems on the assumption that the set of participating users, hardware, and software is highly dynamic and unpredictable. Clear invariants that govern the entire system's execution should exist. (Kindberg & Fox, 2002, p. 71)*

When speaking of adaptations to the environment and context in mobile systems it is usually the mobile computer itself that adapts and reacts, but in the era of ubiquitous computing all the devices have to adapt to each other. The device entering an environment has to enhance the devices currently there, and the devices currently there have to enhance the device entering - they help each other complete tasks and service the users (Buxton, 2013).

As ubiquitous computing devices need to adapt to changes in their environments and contexts, they need to be able to sense these changes. From this need, the area of context-aware computing rose.

3.1.2 Context-aware computing

Knowing the system's and the user's context is essential to ubiquitous computing. Even though context is most often associated with location (Chen & Kotz, 2000), other factors such as what the user is doing and why can have an impact on the context.

> *Context-aware computing is the ability of a mobile user's applications to discover and react to changes in the environment they are situated in. (Schilit & Theimer, 1994, p. 23)*

Because of the prevalent use of location in context-aware systems, they are sometimes referred to as location-sensitive, location-aware or location based systems. Abowd and Mynatt (2000, p. 37) propose the use of the five W's in the area of context-aware systems, stating that a context is based on more than simply the user's location:

- Who: Information about the user and other people in the user's environment may affect the user's actions and activities.
- What: The actions and activities the user is performing.
- Where: The location of the user or system, which is the aspect of context that is used most often by context-aware systems.
- When: The time of day, day of the week, and how long a user has been at a specific location are among the ways the passage of time can affect the user's context.
- Why: Giving a reason to the "What", which sometimes can be accomplished through the use of health sensors such as heart rate or accelerometer.

Chen and Kotz (2000) defined two types of context-aware systems, encompassing different degrees of user control:

1. Active context-aware systems: where the application automatically adapts to the newly sensed context, changing its behavior accordingly.
2. Passive context-aware systems: where the application presents the newly sensed context to the user or stores it for the user to retrieve later.

Using Chen and Kotz' distinction of context-aware systems, Barkhuus and Dey (2003) found that even though users felt more loss of control the more automated their systems were, they preferred using automated systems "as long as the application's usefulness is greater than the cost of limited control". In order to grant the users more control, or as much control as each individual user would like, Hardian, Indulska, and Henricksen (2006) advocates the use of what they call "preference-based decision support". Here, users can personalize their systems and tell them how to behave in certain contexts in the future. For the users to be able to control this, however, the systems need to reveal the aspects of the context that motivate their current behavior. Page, Johnsgard, Albert, and Allen (1996) found that users often customize the functionality of systems, but not so often customize the appearance of the interface. They found no correlation between how much a user customized the software and their experience with computers, their occupation, gender, age or education. However, when the user used the software more, he customized it more. Barkhuus and Dey argue that users are more prone to customize their phones as they are more personal in nature than a desktop computer.

Bellotti and Edwards (2001) suggest that as long as there's a slight doubt in what the user's desired outcome of a situation is, the system must provide an effective way to correct the actions of the system. If, however, there is significant doubt, the user must be able to confirm the action before the system takes it. They suggest this even though they bring forward the issue of users being bombarded and bothered by requests and prompts from the system. They mean that an autonomous system executing the wrong actions are seen as more bothersome than one needing the user to make a final call. A survey conducted by ResearchNow on 1,000 consumers, however, showed that push notifications not giving relevant information, not providing enough value and being too "annoying" were the main reasons users turned them off (GSMA, 2014).

In the next section I will explain a term related to context-aware and ubiquitous computing called calm computing, where systems glide into the background and demand little user attention.

3.1.3 Calm computing

Technology that is simply used as an aid when completing other tasks, instead of as the main task itself, and that still lets you focus on the task at hand, is called calm technology. It is in the periphery, it is unobtrusive and does what it's supposed to do - nothing more, nothing less. Additionally, it moves smoothly and easily from the periphery to the center of our attention when needed, and then back again to the periphery (Weiser & Brown, 1996). We only call for its attention, or it for ours - i.e. moving it to the center of our attention - when needed.

> *We use 'periphery' to name what we are attuned to without attending to explicitly. Ordinarily when driving our attention is centered on the road, the radio, our passenger, but not the noise of the engine. But an unusual noise is noticed immediately, showing that we were attuned to the noise in the periphery, and could come quickly to attend to it. (Weiser & Brown, 1996)*

A substantial challenge in calm computing and technology is finding the balance between simplicity and quantity of functionalities (Weiser, Gold, & Brown, 1999). Systems that have more functionality tend to be more complex, and hence less calm - they can more easily intrude and grab your attention even when it's not needed.

A system can be varying degrees of calm, it is not a binary "calm" or "not calm". Even though the characteristic of calm devices and systems are important for ubiquitous computing, one can argue that when moving into the era of ubiquitous computing one should not strive to make all systems a perfectly calm. Kindberg and Fox (2002) argues that one should establish a clear "semantic Rubicon", i.e. a division between what the system should do and be responsible for and what the user should do and be responsible for when it comes to higher-level decision making. For instance, if the system has difficulties figuring out whether devices in the new environment are trustworthy, or if adapting would be useful at all to the user, the user can step in and make the decision himself. Kindberg and Fox say that the users should be able to step in when ambiguity arises and resolve the situation. Chen and Kotz (2000) state that active context-aware systems, where the systems automatically adapt without consulting the user, will help make systems as calm as possible.

Regardless of the benefits of ubiquitous, context-aware and calm computing, the fact that all these fields want the computer to slide into the background raises some concerns about privacy, which will be handled in the next section.

3.1.4 Privacy concerns

Not soon after the term ubiquitous computing was coined, people started raising concerns. Even though the implementation of this new wave of computing and its accompanying infrastructure was thought to be "years or decades away" (Araya, 1995, p. 4), the concern for invasion of privacy quickly started to get discussed. Gudymenko and Borcea-Pfitzmann (2011) define privacy as follows:

> *Privacy of an entity is the result of negotiating and enforcing when, how, to what extent, and in which context which data of this entity is disclosed to whom. (Gudymenko & Borcea-Pfitzmann, 2011)*

Computing devices surrounding us throughout the environment will know about our location, activities, habits and patterns. Spiekermann (2005) mentions two distinct reasons for the loss of privacy in the context of ubiquitous computing: (1) People losing control over being accessed, and (2) Lack of control over information use and maintenance once people or their objects have been accessed. The first point is backed up by Abowd and Mynatt (2000) stating that users find the "invisible" notion of wireless communication fearful and off-putting as it doesn't allow them to control or to know when information is being exchanged - it happens "behind their backs". Abowd and Mynatt state that being able to know when one is being surveilled is just as important in the technological space as in the physical space. Before one can control or completely stop the surveillance one has to be aware that it is taking place. According to Langheinrich, Coroama, Bohn, and Rohs (2002) a complete loss of privacy is the most direct fear associated with ubiquitous computing, and they as well as Röcker (2010) and Lucky (1999) mention the fear of the saying "if these walls could talk" coming true.

The second point Spiekermann (2005) brings forward concerns unauthorized secondary use of the gathered data. Friedewald and Raabe (2011) mention that the use of electronic tickets such as RFID bus or ski passes can help collect detailed movement and behavior patterns, potentially used for marketing or other areas, even though this pattern collection might not be known by the user. As Spiekermann puts it, however:

> *This secondary use (and abuse) of information is not possible if there has not been access in the first place. This implies that controlling access is a crucial part of the privacy equation in [ubiquitous computing]. (Spiekermann, 2005, p. 3)*

Despite the critique and concerns that accompany ubiquitous computing, however, Abowd and Mynatt (2000) do not see them as reason enough to stop all research and development in the area, but rather to work on solutions that address the concerns. As a result of further research in the area, computing devices started getting implemented into home appliances and other previously non-electronic entities. The "Internet of Things" was born.

3.1.5 Internet of Things

In 1999, staff at the MIT Auto-Id Center wanted to be able to digitally identify non-electronic entities, or "things", in order to ship and track goods more easily. They coined the concept "Internet of Things" (IoT). Even though the main focus of the Auto-Id center staff was identification, IoT quickly evolved to also describing the information of these "things" and their relationships (Kuniavsky, 2010, pp. 78-80).

With the growing discussion and desire to connect devices to the Internet, the term "Internet of Things" (IoT) rose to popularity in the early 2000's. Ever since, the world has moved closer to ubiquitous computing becoming a reality. Computing devices surround us and interoperate, albeit not always spontaneously or without the help of human interaction. The main form of communication over the Internet has for a long time been human-human, where humans use the Internet and its technologies and connected devices to communicate information to each other. Now, however, more and more communication is on a human-thing and thing-thing basis (Tan & Wang, 2010). These "things" are any devices that are in some way connected to the Internet, be it a watch, a car, a TV, a dishwasher, a chair, a pen etc.. Weiser and Seely Brown, while working on the beginnings of ubiquitous computing, started connecting "things" (more specifically a coffee pot) to the Ethernet and broadcasting its state as early as the late 1980's. As Seely Brown put it when in 2014 he reflected on his and Weiser's work:

> *It was the Internet of Things before the Internet of Things! The coffee machine would announce on the Ethernet when a fresh pot of coffee was ready. (re:form, 2014, p. 10)*

However, a coke machine at the Carnegie Mellon University was connected to the Internet as early as in 1982. You could check the availability of products, as well as checking whether what you wanted was cold or not as long as you were connected to the Internet. In 2015 these "things" and devices are more widespread than ever as Smart TVs, RFID bus passes, lighting controlled from your phone and similar things are a part of the everyday life of the layman. These "things" are getting more and more prevalent throughout society - they are becoming ubiquitous. The different things and devices need different identities and ways to communicate, spawning the innovation and use of several different technologies.

The next section elaborates on how computers can sense and "talk to" the physical non-electronic world, and how the Internet of Things can be continued.

3.2 Connecting computers to the physical world

The need for computers to be integrated and communicate with the physical non-electronic world grows as we transition to the world of ubiquitous computing. The technologies this section details were discussed during my studies, both in the interviews and the focus group.

3.2.1 Magnetic stripes and smart cards

The magnetic stripe card, invented in the late 1960s at IBM, allowed for a quick and easy way to store information for e.g. identification or payment. Before it, the process of paying with a credit card or checking in on an airplane was too cumbersome and took too long, and as a solution the magnetic stripe card was invented, making information available from a small physical container and able to be retrieved quickly (Svigals, 2012).

In the middle of the 1980s the smart cards started emerging. These cards are credit cards just like the magnetic striped ones, however, they have a microprocessor embedded in them. Outside Canada and the United States these smart cards grew to be the most popular, although often accompanied by a magnetic stripe on the same card. These smart cards offer more functionality than a magnetic striped card, such as security features which enable greater levels of encryption.

Making a purchase with a magnetic stripe card, the user swipes the stripe through a reader and enters a PIN code On the smart cards, the swiping motion is replaced with sticking the part of the card containing the chip into a reader, then entering a PIN. These technologies enabled quick identification and verification of information, as opposed to earlier methods.

3.2.2 QR codes

Quick Response (QR) codes were invented in 1994 by Denso Wave as a 2D barcode, being able to store more information than the "classic" 1D barcodes and being scannable with a relatively high speed. A mobile application can use the phone's camera in order to scan and interpret the code, which in the use of mobile applications and advertisement boards often decodes into a website URL or an e-mail address (Furht, 2011, p. 341) Figure 3.1 shows a popularly used square QR code on the left, compared to a 1D barcode on the right. A QR code consists of several special fields to allow the scanner to orient the code correctly and potentially correct any errors.

Figure 3.1: A QR code (left) and a 1D barcode (right)

In order to read the information stored in a QR code, the user has to have a device capable of scanning and decoding the information. Different QR scanner applications can be downloaded on smartphones, however, they are usually not pre-installed when purchasing a new phone. As a result, users have to seek out these applications themselves. On a smartphone the QR scanner uses the phones camera to capture the QR code, then proceeds to decode the information, presenting it to the user when completed. Some QR codes are printed small, some large, and the camera can have trouble focusing on the QR code to read it properly. As the user first has to find the QR scanner application and focus his camera on the code, retrieving human-readable information from a QR code can take multiple seconds. As long as the camera can get a good focus on the QR code, it can be decoded, meaning that they work at different distances and sizes.

3.2.3 RFID and NFC

During the late 1990s and early 2000s many of the ubiquitous computing systems were based on RFID chips (Langheinrich et al., 2002), (Sakamura, 2003), (Friedewald & Raabe, 2011). Radio frequency identification (RFID) is a way of identifying devices and objects through the use of radio waves. A reader picks up that a device is in its proximity and is transferred the ID of said device. The reader then executes a set of operations based on the ID of the device, such as logging that the device was within the readers proximity, or contacting a server to authorize access. There are two different kinds of RFID devices: the active and the passive (Want, 2006). The active devices need to be supplied power e.g. through a battery or a connected to a powered infrastructure. Passive devices do not requires batteries, hence they can be smaller, lighter, often cost less and their lifetime is not as limited. The device reading the passive RFID device is in charge of providing it with power through use of magnetic induction, as is the case of many RFID bus passes, or electromagnetic wave capture.

Magnetic induction enables the use of near-field RFID as the device needs to be close to the power source to be enabled. These devices then send data to the reader (power enabler) by producing its own magnetic field and changing it over time. This will in turn change the readers magnetic field and current, allowing the reader to recover the information sent. This is a technique known as load modulation. The more bits needed for expressing a device's ID, the higher the frequency of operation needs to be and the read time increases. Additionally, the amount of energy transferred rapidly drops the farther away from the reader the device gets.

NFC (Near Field Communication) is a specification building on the near-field RFID technology, extending it by adding security measures and a common data format (STMicroelectronics, 2015). NFC is what is normally used in bus passes, access cards, debit cards, credit cards and phones. The NFC chip in the device that belongs to the user transfer information to a reader, such as a payment terminal or scanner by a door. The user simply holds the device or "thing" with the NFC chip at most a few centimeters from the reader, and within milliseconds the information is transferred. Hence, the user has to make a conscious choice as to whether or not he wants his information to be read, just as with the use of QR codes.

3.2.4 Bluetooth beacons

Bluetooth beacons are based on a Bluetooth technology released in 2010 called Bluetooth low energy (BLE), which does not replace classic Bluetooth - they are implemented fairly differently and thus have different use cases. A beacon constantly broadcasts information to any device in its environment that wants to read it. As opposed to QR and NFC, information transfer by the means of Bluetooth beacons does not necessarily require any human interaction, sacrificing user control for autonomy. These following sections describing Bluetooth and beacons are a fair bit more detailed and technical than those describing QR and NFC. As the focus of this thesis is the use of beacons, I would argue that a more technical and detailed elaboration of beacons is appropriate.

Classic Bluetooth

Classic Bluetooth is a wireless technology standard for transmitting data. It uses short-wavelength ultra-high frequency radio waves in the range from 2402 MHz to 2480 MHz, or 2400 MHz to 2483.5 MHz when including guard bands - unused parts of the radio spectrum to avoid or diminish interference. This radio frequency range is part of the 2.4 GHz to 2.5 GHz ISM band. ISM bands are radio bands not used for telecommunications, but rather used for industrial (I), scientific (S), and medical (M) purposes in order to prevent interference on and from other radio communication. In order to reduce the interference on signals, Bluetooth uses a technique called Frequency Hopping Spread Spectrum (FHSS). FHSS spreads the signal over a range of frequencies, rapidly and frequently changing ("hopping") which frequency its currently transmitting at. Additionally, spread spectrum signals are more difficult to intercept and sniff. Bluetooth makes use of 79 different channels, each with 1 MHz of bandwidth (SIG, 2010, Vol. 1 pp. 18-19).

Bluetooth low energy

In 2010 the Bluetooth 4.0 specification was released, which included the new technology of Bluetooth low energy (BLE). This new BLE technology differs from classic Bluetooth in several ways, not making it an improvement or upgrade of classic Bluetooth, but a new technology alongside classic Bluetooth. A device can support either BLE, classic Bluetooth or both (Torvmark, 2014). Where classic Bluetooth makes use of 79 different 1 MHz channels, BLE makes use of 40 different 2 MHz channels (SIG, 2010, p. 126). And where classic Bluetooth uses FHSS as the spread spectrum technique, BLE uses Direct-Sequence Spread Spectrum (DSSS) ('Bluetooth Low Energy Regulatory Aspects', 2011, p. 12). In each of these 40 BLE channels, data is transmitted using Gaussian frequency shift-key (GFSK) modulation. Modulation is the process of modifying or varying a radio signal in order to encode it with data. As GFSK is a frequency modulation technique, it encodes data by introducing variation in frequency to the signal. Of the 40 BLE channels, only 3 are used for broadcasting and the remaining 37 are used for data transmission between paired devices. These three broadcasting channels were chosen in an attempt to avoid the main Wi-Fi channels as best as possible in order to further reduce interference. When broadcasting, the BLE device transmits an identical packet over the three channels sequentially, one after the other (Vlugt, 2013).

Beacons

Beacons, in its traditional sense, have been used since long before the introduction of electronics. A beacon in its simplest form is something that is designed specifically to attract attention to itself or its location. The semantics and meaning of each beacon may be different, depending on the context in which it is placed and what information it broadcasts. Hence, a beacon may take the form of a cairn on the top of a mountain or by the sea, of a lighthouse to guide boats to safety, of a lit fire to warn about incoming enemy troops, of a flashing light indicating an ambulance needing to get through traffic, and so on.

In the context of wireless technologies beacons are devices broadcasting or advertising small pieces of information which allow everyone within the range of the beacon to pick up said information (Lindh, 2015). As beacons make use of BLE, branded as Bluetooth Smart, they can often broadcast for years on a single battery, depending on broadcasting frequency, signal strength and whether or not the beacon is connectable. Beacons can also be powered by a DC power supply through e.g. USB, or through different power harvesting methods such a solar power or mechanical pressure. BLE devices manage to last for such a long time on such small power supplies because they have a very short transmission time of small packets, and between transmissions they go into sleep mode, conserving battery. BLE devices can estimate their outdoor or indoor position based on the signal they receive from nearby beacons, which is elaborated on later in 3.2.4.

There are four different operation states for a BLE device. In the different states the BLE device steps into different roles, and in the first two roles the operation is connection-based or connectable, meaning it supports two-way communication. The different roles enable the BLE devices to be used in a range of different applications (Lindh, 2015).

- Peripheral
  A peripheral device can operate as a slave once a connection to another device is established. The peripheral device broadcasts information. Examples of peripheral BLE devices are health thermometers or heart rate sensors.

- Central
  A central device can operate as a master with an established connection to one or more devices at the same time. The central device scans for broadcasters and it is the one initiating the connection. Examples of central BLE devices are smartphones or computers.

The other two roles of operation are non-connectable, meaning they only support one-way communication, i.e. classic broadcasting or information gathering.

- Broadcast
  A broadcast device broadcasts information without any knowledge of the devices nearby or any established connection. Examples of broadcast BLE devices are temperature sensors broadcasting the current temperature or a tag for tracking the location of a packet broadcasting its ID.

- Observer
  An observer device scans for broadcasts from other devices. It collects data and might process it, such as displaying the data on a screen. The observer can not talk back to the broadcasting devices because it can not establish a connection, hence the observer is non-connectable. It receives data, it does not send it. An example of an observer BLE device is a display presenting temperature data broadcast from a thermometer or an oven.

A beacon, being a broadcasting device, fits into the roles of peripheral (connectable) and broadcast (non-connectable) devices. The connectable beacons are more configurable and controllable than the non-connectable beacons, making them useful in different contexts. As the broadcasting beacons are non-connectable they can achieve the lowest possible power consumption by simply transmitting and going back to sleep mode, transmitting and going back to sleep mode, and so on. The peripheral beacons, however, are connectable and might use power on receiving and processing information from a connected device, giving it more to do than simply transmitting and sleeping. Hence, the broadcasting beacons have the lowest power consumption and the longest battery life given the same supply.

BLE packets

Appendix A explains the fields a BLE transmission packet consists of and what they add. In 2013, Apple introduced a BLE technology called iBeacon, building on the standard BLE packet. This widely used protocol adds the fields UUID, Major and Minor to the standard packet. These fields are used as identifiers by applications reading iBeacon packets where the UUID represents the application and use case, while Major and Minor are used for sub divisioning and further detailing what a specific beacon is used for (Apple, 2014). This sub divisioning can help optimize the scanning for beacons as the application can discard the rest of the packet if the UUID or the Major does not match what the application is looking for. On the application software side, the iBeacon framework provides an easy way for developers to read the distance from the reading device to the beacon, however, this information is only given as "Immediate", "Near", "Far" and "Unknown", where "Near" is 1-3 meters in clear line of sight.

In 2015 Google released their own beacon protocol called Eddystone. This protocol encompasses three different packet types: UID, URL and TLM (Google, 2015). The UID packet is similar to the iBeacon protocol. The URL packet transmits a compressed URL in order to connect everyday "things" and the Internet to the physical world. A parking meter, for instance, can broadcast its URL to allow users to pay for their parking space through the web. The TLM, or telemetry, packet is used to monitor the state of the beacon, and can only be used as an add on to the UID or URL packets, not by itself.

Both iBeacon and Eddystone urge deployers of beacons to measure each devices signal strength at a specified distance (0 meters for Eddystone, 1 meter for iBeacon), and encode this into the packet the beacon broadcasts. This value, often referred to as TxPower, is used by the application reading the BLE packet in order to estimate its distance from the beacon, as explained in section 3.2.4.

Power consumption

The battery of the beacons can last for anything from a few weeks to several years, depending on the device hardware, battery, broadcasting rate and packet size. Some beacons broadcast no more than their ID, while others broadcast a message or values from sensors. The broadcasting rate is a significant factor on the battery life of the beacon, where increasing the rate will increase the power consumption. However, the higher the broadcast rate, the faster an application can discover the beacon and the more accurate and quickly the distance estimation can become (GSMA, 2014, p. 31). The trade-off between power consumption and responsiveness is something that needs to be discussed on a case by case basis. In some cases, however, a DC power supply can be used, eliminating the problem of power consumption. A non-connectable BLE beacon can not have a broadcasting interval less than 100 ms, meaning it can not broadcast any faster than once every 100 ms (SIG, 2014, p. 66). As a beacon is a broadcasting device, only the three broadcasting channels are used by both transmitting and scanning devices, greatly reducing the power consumption over classic Bluetooth.

Just as the act of broadcasting consumes power for the beacons, scanning for beacons consume power for smartphones or other BLE enabled devices. Different strategies can be implemented in order to lower the power consumption, but just as in the broadcasting rate of the beacons, the scanning interval will affect the responsiveness of the application. GSMA (2014, p. 41) gives examples of four different strategies:

- Only scanning during the opening hours of the stores the app relates to
- Only scanning when the consumer is not stationary
- Only scanning when within a larger geo-fence containing the beacon location
- Only scanning for a couple of seconds, then going back to sleep for about 20 seconds if no beacon is found

A study done by Aislelabs (2014) shows that the amount of beacons placed in the environment can affect the smartphone's power consumption. During the study, they let different phones scan for BLE devices continuously for an hour with an interval of 1 second, meaning the phone scans for 1 second, pauses for 1 second, scans for 1 second, and so on. As can be seen in Figure 3.2 the Android phones tested consumed less power than the iOS phones for the same amount of beacons placed in their immediate surroundings.

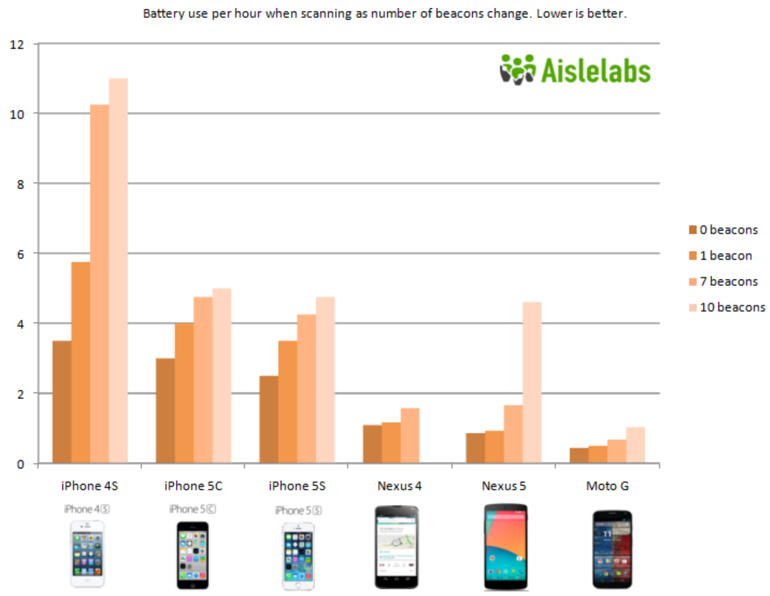

Figure 3.2: Battery consumption after a one hour BLE scan with a 1 second scan interval (Aislelabs, 2014)

The study also showed that the power consumption of BLE scans are improving dramatically with the release of new phones and technology. With 10 beacons near the phone, the iPhone 4S consumed 14.50 % battery while scanning, as opposed to a baseline of 3.50 % while not scanning. In the same situation the iPhone 5S consumed 7 % battery while scanning, as opposed to a baseline of 2.25 % (Aislelabs, 2014). According to Aislelabs, however, this low power consumption of the Moto G Android phone might be the cause of the Moto G only processing about 30 % of the BLE packets. If this is the case, this can have an effect on the responsiveness of the applications.

Interference and propagation

As Bluetooth is based on radio technology, it is prone to loss or disturbance of signal through interference (HP, 2002). Radio waves can superpose, meaning they combine into one wave with properties reflecting the properties of the two waves before superposing. Given that Bluetooth resides in the 2.4 GHz ISM band together with other prevalent technologies such as Wi-Fi and microwave ovens, interference is a legitimate concern. As discussed earlier, BLE uses three broadcasting channels specifically chosen to try to avoid interference from the Wi-Fi channels accompanied by the use of a spread-spectrum technique. Another measure that is implemented for ease of co-existence is a pseudo-random delay to the broadcasting interval between 0 and 10 ms as an attempt to avoid interference of beacons broadcasting at the exact same time (SIG, 2014, p. 66).

In addition to the problem of signal interference, the signal can get weaker as it propagates through space, some mediums affecting the signal more than others. This phenomenon is called attenuation, and may drastically reduce the signal strength and the reach of the beacon. For instance, if a beacon is placed on a metal shelf or within a metal box the level of attenuation can be so strong that the ability to properly calculate distance to it becomes severely inaccurate (Vlugt, 2013, p. 9). Interference and attenuation can affect the accuracy of location estimation based on BLE beacons.

Location estimation

In the field of wireless technologies, the received signal strength (RSS) can be used to estimate the distance from the emitting device. The RSS is often represented by a number called the Received Signal Strength Indicator (RSSI). Bahl and Padmanabhan (2000) studied the use of in-building location tracking where a system tracked the location of different devices within the building using wireless local area networks (WLANs). A very similar system (LEASE) was studied by Krishnan, Krishnakumar, Ju, Mallows, and Gamt (2004). Roos et al. (2002) studied indoor tracking where the users "mobile terminal" performed the signal strength measurements from several WLAN transmitters placed throughout the room. Subhan et al. (2011), amongst others, used a technique called fingerprinting to track indoor position where the RSSI of different locations throughout the area are a-priori logged and mapped during what is often referred to as the "offline phase". This mapping is called the RSSI "fingerprint" of the area. The "online phase" then uses this fingerprint to look up the location of the device, often by sending a request to a server. Having the fingerprints stored on a server instead of locally on each calculating device allows the developers to more easily change the layout and the fingerprints of the area. The estimation can also be conducted without the use of an offline phase, but this requires a mathematical model of the signal propagation. According to Salas (2014) who researched location tracking on BLE beacons specifically, using a propagation model called the log-distance path loss model is appropriate for indoor environments, granting him a location estimation of 1 to 1.5 meter accuracy while in clear line of sight of a beacon.

Estimating the distance from a single beacon will result in a fairly inaccurate location estimation as the user can be placed anywhere on the circle with a radius equal to the calculated distance to the beacon. Using a technique called trilateration will help narrow down the possible points of the circle the user can be located. Trilateration uses the distance measurement from several beacons in order to generate several of these circles and then calculate the point at which these circles intersect. Figure 3.3 shows an illustration of the concept. B. Cook, Buckberry, Scowcroft, Mitchell, and Allen (2005) used trilateration in order to more accurately estimate a users indoor position with the use of WiFi signals. Jr (2007) used trilateration to localize mining equipment in three dimensional space.

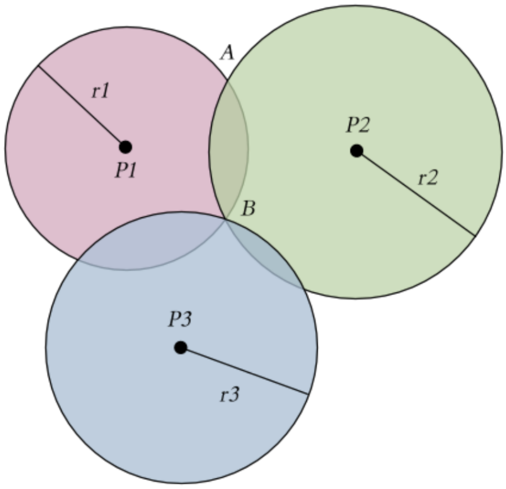

Figure 3.3: Illustrating how the use of more beacons can increase the accuracy of the location estimation (Salas, 2014)

As radio signals tend to fluctuate, especially in indoor environments, the readings can vary substantially even though the user has not moved. To combat this issue B. Cook et al. (2005) suggests averaging the RSSI before calculating the distance to the beacon, or averaging the estimated position. The more accurate the average the more readings are required, hence the longer it will take to calculate a position estimate.

The next section describes a less technical aspect of Bluetooth beacons - the privacy concerns that they introduce.

Beacon privacy

A major privacy problem in the field of radio based technologies is that the communication and interaction is invisible to the naked human eye, meaning the users never know when communication takes place (Want, 2006). As the BLE broadcasting packets are inherently open, any scanning device can read the contained information and malicious users can then replicate the signal either through a phone or by implementing it in their own beacon. What this new malicious beacon ends up doing to the application can vary depending on the application itself, ranging from preventing accurate location estimation to spamming the user with notifications or having the application execute automatic actions at inappropriate times, such as automatic payment (Vlugt, 2013, p. 9).

The Norwegian Data Protection Authority saw in 2014 an increase in the use of beacons and Wi-Fi as a means of tracking people's movements through both public and private space. The amount of time this information is stored can range from a few minutes to indefinitely. The Norwegian Data Protection Authority sees this as a challenge for anonymous passage and for being in control of your own personal data, as appropriate consent is often not obtained (Datatilsynet, 2014, p. 51). With the use of Eddystone and iBeacon followed by a possible integration of BLE scanning in newer versions of Android, the issue of appropriate consent can get even more troubling. The threshold for beacon scanning is lowered as the user won't necessarily have to download a specific application to their smartphone - scanning for and processing Eddystone or iBeacons can become a built-in functionality of the operating system. Eddystone scanning already becoming an integrated part of the Chrome web browser for both Android and iOS[1]. The Norwegian Data Protection Authority is concerned that disabling Bluetooth on your phone might be an inadequate countermeasure, as the phone can by itself re-enable Bluetooth without displaying the Bluetooth icon in the task bar, giving no indication to the user that Bluetooth indeed is enabled (Årnes, 2016).

According to the Norwegian Data Protection Authority as well as Norwegian law, informed consent is required in order to collect, treat and process personal information, and the consent needs to be "voluntary, explicit and informed". A voluntarily provided consent means that it is not given as a result of being forced or threatened. An explicitly provided consent means that it should be unambiguous that the person indeed has provided his or her consent, and it needs to happen through an active task:

[1]http://blog.chromium.org/2016/02/the-physical-web-expands-to-chrome-for_10.html

> *For instance, it is not sufficient to provide information about the processing of personal information at a web page, followed by a sentence which states that the user consents by using the system. (Datatilsynet, 2015, p. 31)*

An informed consent means that the person should be informed as to what the intended use of the information is, who is responsible for it, which kind of information is collected and so on. Additionally, consent should be able to be retracted and the system should facilitate this. The information should not be stored for a longer period of time than what is necessary and the processing of information should not go beyond what is necessary based on the purpose of the data collection (Datatilsynet, 2015, pp. 31-32). This supports what Spiekermann (2005) says about unsolicited and unknown secondary use of the user's data. The system should not use gathered data for purposes beyond what the user has been informed about and has confirmed.

Gudymenko and Borcea-Pfitzmann (2011) argue that having information only flow from the infrastructure to the user, not the other way around, will enhance the user privacy regarding ubiquitous computing. This is the case when dealing with beacons, as the beacons broadcast their information to the users' smartphones and each user's smartphone calculates its location based on this information. However, the smartphone can then in turn send this context data to a server which stores contextual information about you in your "profile". Hence, beacons and this "infrastructure-to-user" communication is no panacea for ubiquitous computing privacy concerns.

3.3 User experience

The term user experience (UX) has been defined in different ways by different researchers and practitioners throughout the years. Kuniavsky (2010, p. 14) says that no definition of user experience is complete without taking into account the entirety of what a user experiences when interacting with a system. The related term usability encompasses the ease and effectiveness of use of a system, but in comparison, the user experience needs to take into account factors such as the user's satisfaction while operating the system, not just objective factors such as average user errors or number of clicks needed to perform an operation. Often, however, usability and user experience are used interchangeably. Kuniavsky's definition of user experience is as follows:

> *The user experience is the totality of end users' perceptions as they interact with a product or service. These perceptions include effectiveness [...], efficiency [...], emotional satisfaction [...], and the quality of the relationship with the entity that created the product or service [...]. (Kuniavsky, 2010, p. 14)*

This is in line with the findings of Law, Roto, Hassenzahl, Vermeeren, and Kort (2009), saying that the user experience has to take into account the subjective nature of experiences and perceived usability. The ISO definition[2] of user experience says that it is "A person's perceptions and responses that result from the use or anticipated use of a product, system or service", meaning a user has to feel satisfied on an emotional as well as instrumental level while interacting with the system in order to find the user experience good - a user's internal, subjective state is involved (Hassenzahl & Tractinsky, 2006). The user's anticipations, predictions and part of their internal state are stored in their "mental model". Before I explain what a mental model is, I will explain what a mental model is not - a conceptual model.

3.3.1 Conceptual models

When designers create a system they typically have a specific workflow or way of using the system in mind. The conceptual model of a system is an abstraction which states how different entities and sub-systems are defined, how they relate, how they inter-operate, and explicitly or implicitly explains how each particular entity should be used to achieve a specific goal. It is a model defining the system and often explaining its intended use, and often decomposing high-level requirements into a more detailed specification. This

[2] ISO 9241-210:2010 Ergonomics of human-system interaction

will help the developers and designers share the same idea and develop towards the same goal and assist them in architectural and functional decision making. Additionally, the visualization of the conceptual model helps communicate the idea of the system to the stakeholders, eliminating ambiguity (Sommerville, 2011, p. 154).

There has been created numerous tools to visualize and express conceptual models, one of which is UML, Unified Modelling Language (James Rumbaugh, 2004), (P. Stevens & Pooley, 2006). Through UML one can visualize a system's different components, their architecture, structure and the interaction between them. UML is built up of several different types of diagrams divided into two categories: structural and behavioral diagrams. Structural diagrams such as class diagrams or object diagrams visualize the internal structure of entities in the system, not necessarily explaining how the entities are related to the rest of the system. Behavioral diagrams, on the other hand, such as activity diagrams or sequence diagrams help express the interaction between different components or the intended workflow between the user and the system (Sommerville, 2011, pp. 119-142). Figure 3.4 shows an example of an activity diagram.

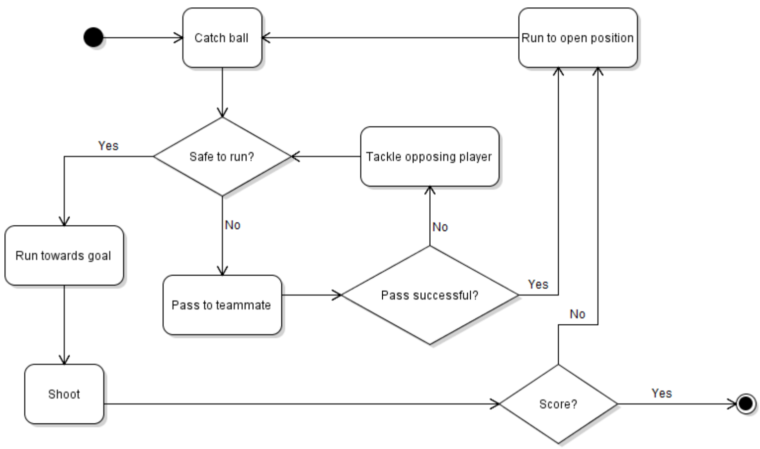

Figure 3.4: Example activity diagram of a football match

Closely related to conceptual models are mental models, however, the mental models focus on how the user understands and perceives the system.

3.3.2 Mental models

A term coined by Kenneth Craik in 1943, a user's mental model of a system is how the user perceives that system and believes is the correct way to interact with it, regardless of whether this belief is correct or not (Craik, 1967).

> *Mental models research is fundamentally concerned with understanding human knowledge about the world. (Gentner & Stevens, 1983, p. 1)*

Each user has his or her own mental model, which is created and evolve through direct or indirect exposure to the system, helping users predict and anticipate the system's behavior (Gentner & Stevens, 1983, p. 12). Direct exposure can be interacting with the system or reading the instruction manual, while indirect exposure can be talking to a friend or coworker about the system and exchanging experiences and understandings, or being exposed to other similar systems. Mental models help users simplify and understand the world that they live in and the systems and environments they are interacting with. A user's predictions and actions are directly affected by and based on their mental model of the system (Nielsen, 2010). The mental model of a user can often differ from the conceptual model the system was designed around, so the users can believe and act like the system works in different way than it does in reality. When the user has a clear and correct mental model, the interface seems to disappear, letting the user concentrate on his work - the system will become an invisible tool to achieve a goal (Shneiderman & Plaisant, 2010, p. 31).

A not too unusual mix up of mental models that can occur in 2016 happens when a user tries his hands on a touch screen tablet for the first time. On the tablet, the keyboard appears as needed, but only through the instruction of the user, e.g. by clicking an input field. If this relationship between actions and appearance of the keyboard is not correctly understood (i.e. the mental model is not correctly established), the user will feel that the keyboard appears at seemingly random times, and not every time he needs it. Additionally, for mouse based systems, the right mouse button tend to open up a menu with more options, while on touch based systems this is often done by pressing a button for longer than a quick tap (long-click). Keyboard and mouse based systems require a different mental model than their touch screen based brothers, and when changing from one type of system to the other users can mix up their mental models and, at least for some time, forget how the system is supposed to be used - touching a screen that doesn't support touch with their fingers, or searching for their mouse while on a touch tablet.

If a user's mental model is false or incorrect, he may end up failing to accomplish their tasks, causing frustration and possibly temporary or permanent abandonment of the system. As such, developers and designers should, whenever possible, design to standards, conventions and user expectations. Buie (1999) says:

> *Standardization facilitates learning and reduces errors by taking advantage of knowledge the users have gained from other products or from other parts of your product. (Buie, 1999)*

When innovating, however, the system needs to be self-explanatory or the new aspects and interaction methods should be presented in a clear and concise manner so that the users can quickly understand and start using the system. The more complex an innovation is, and the less it can be related to anything a user knows, the greater the chance the user will have of establishing a wrong mental model. This can serve as a challenge when designing for ubiquitous computing systems, as it is a fairly innovative field.

A user's mental model is continuously evolving through interaction, being modified to fit the user's experiences. However, the models are constrained by factors like experiences with similar systems and technical background. Hence, Kuniavsky (2010, p. 274)) writes that when developing and designing new technologies, building on and extending familiar experiences are of importance, as it can help ease the transition. Kuniavsky also notes that when innovating technology, if the user has to adjust to multiple new and innovating functionalities, confusion and misunderstandings might occur. Thus he suggests focusing on what he calls core functionality. Additionally, based on the "confirmation bias" theory (Nickerson, 1998), people can tend to only seek out or remember facts that support their already existing mental models and theories, interpret ambiguous information in a way that supports their model, and completely disregard information that may challenge it.

The term "double-loop learning" was coined by Argyris, Putnam, and Smith (1985, pp. 80-102) in order to explain learning that was grounded in a change in a person's fundamental understanding - their mental model. Single-loop learning, on the other hand, the counterpart of double-loop learning, only results in a temporary change in action, where the person sees that their actions did not result in the desired outcome and they try a new set of actions to achieve the same desired outcome. In double-loop learning the person reflects over his mental model, trying to see how he might have misunderstood the problem. It helps the person ask questions about his own understanding, reasoning and motivation behind his actions, recognizing that a source of the problem may be the way he understands it. However, double-loop learning does demand substantially more energy and focus than what single-loop learning does. Figure 3.5 illustrates the difference between single-loop and double-loop learning.

Figure 3.5: Illustrating single-loop learning as the inner loop, and double-loop learning as the outer loop (Richardson, 2014)

A user's mental model can consist of information that the user is uncertain about and the degree of this uncertainty, but the information is still included because it "seems to work". When a model consists of such recognized uncertainties a user will normally operate with greater caution. These uncertainties are often the result of getting mental models of different yet similar systems mixed up:

> *A person's mental model can include knowledge or beliefs that are thought to be of a doubtful validity. [...]These doubts and superstitions govern behavior and enforce extra caution when performing operations. This is especially apt to be the case when a person has experience with a number of different systems, all very similar, but each with some slightly different set of operating principles. (Gentner & Stevens, 1983, p. 8)*

In short, the difference between a mental model and a conceptual model is that the former is individual to each user and represents their understanding of the system, while the latter explains how the system is intended to work:

> *Conceptual models are devised as tools for the understanding or teaching of physical systems. Mental models are what people really have in their heads and what guides their use of things. Ideally, there ought to be a direct and simple relationship between the conceptual and the mental model. All too often, however, this is not the case. (Gentner & Stevens, 1983, p. 12)*

When the mismatch between the conceptual model and a user's mental model is big enough, the user will stumble upon failure after failure, possibly end up frustrated and abandon the system. When the system behaves, to the user, in seemingly strange and random ways, the user might feel a loss of control. Regardless of the actual control a user has in a system, the control a user perceives that he has is integral to his satisfaction of the system as a whole. This is a key part of the research in this thesis and I will come back to this during the discussion.

3.3.3 Perception of control

One of Shneiderman and Plaisant (2010, p. 89)'s eight golden rules of interface design is to support an internal locus of control. They argue that users desire a sense of control, that they are "in charge of the interface", that their actions directly affect the interface, and that the familiar behavior of the system is not broken by surprises.

> *Users' desire for mastery and a sense of accomplishment can be undermined by an overly enthusiastic interface that does more than they want. (Shneiderman & Plaisant, 2010, p. 592)*

If the users don't feel like they are in control, Barkhuus and Dey (2003) say, frustration can arise and they can be more prone to turn away from the system or service temporarily or permanently. Here a connection can be drawn to users abandoning a system if their mental model is substantially different from how the system is intended to be used, causing unintended behavior - the users will lose their sense of control. If people perceive they have control they will feel more motivated to be in the situation or use the technology (Spiekermann, 2005). Ajzen (1991) argue that a person perceiving control over a situation is more prone to approach it, and one perceiving less control is less likely to approach it. The important aspect here is not how in control the users actually are, but how in control they feel and perceive they are. A user can be motivated and satisfied while using a system where he perceives and believes that he is in control, even though he in actuality has little or no control (Skinner, 1996).

According to van der Heijden (2003), users tend to feel more anxious the less control they perceive, however, the desired amount of control is different from person to person. Kuniavsky (2010) says that making systems more automated and sacrificing control can be tempting in order to increase efficiency, however, the users may still value some control:

> *Grounding new technological experiences in how people define control, where they value it, and when to preserve it is one of the great challenges of ubiquitous computing user experience design (Kuniavsky, 2010, p. 260)*

Dourish (2004) argues that ubiquitous computing is strongly related to tangible computing - computers integrated in tangible objects. These are systems in which the physical world and the computing world are integrated in some way, giving the physical objects some virtual or digital features. A good example of a tangible computing system is the Marble Answering Machine designed by Durrell Bishop in 1992. In this system, a message on an answering machine is represented as a marble laying in a bowl. The user can then pick up this marble and drop it in a hole on the machine in order to listen to the message (Poynor, 1995). According to Dourish there are often no "single point of interaction" when dealing with tangible computers or ubiquitous systems. Many smaller devices situated in the environment come together to accomplish a task, forcing the user to attend to and possibly control several devices at once. The activities on these different devices need to be done in coordination, and the sequential order of the home computer no longer holds, which can make the users feel at a loss of control.

According to Spiekermann (2008) people tend to resist when automation is introduced, feeling that the loss of control becomes too big, and often not seeing the point of automating the task. When losing too much control, people tend to feel helpless. This especially holds true when talking about consumer technology, as there is not as much resistance towards automation in the workplace. In her study, Spiekermann found that users that perceived the most control over a system also had the most positive attitude towards it.

Weiser et al. (1999) sees the problem of user control in ubiquitous computing as an important one:

> *Maintaining simplicity and control simultaneously is still one of the major open questions facing ubiquitous computing research. (Weiser et al., 1999)*

When automating systems, they are often made more simple, at least in interaction. However, this comes at the cost of user control. To find a good balance here is important. To summarize, the users' perception of control is integral to their willingness to use a system and their attitudes toward it, making this an area of concern when automating systems through ubiquitous computing. As it is an important aspect of my research, I will revisit this topic during the thesis' discussion.

Chapter 4

Methods

This chapter describes the methods used in the thesis to answer the research questions. A fairly basic Android application with limited functionality was developed as a prototype. This prototype was then tested with 7 users in order to get the users' feedback and reactions. During and after these user tests the participants were interviewed semi-structurally. Lastly, a focus group with 4 participants, none of whom was part of the user tests, was conducted based on the research questions and topics encountered during user testing and interviewing. All user tests, interviews and the focus group were recorded on audio, transcribed and reviewed through a technique called open coding. The first three user tests were carried out in a large classroom at the University of Oslo, and the subsequent four user tests were carried out in-situ on the metro in Oslo. The focus group with four participants was conducted in a classroom.

4.1 Paradigm and methodology

Where quantitative research is more rooted in mathematics, numbers and statistics, qualitative research focuses on the social and cultural aspects through the use of in-depth techniques like interviews and observations. Based on this, quantitative research is often referred to as "hard", while qualitative research is often referred to as "soft". The participants' meanings and viewpoints, as well as the social context are often lost with quantitative reasearch. There are several different paradigms in both these areas of research. M. D. Myers (1997) talks about three paradigms:

1. Positivist paradigm: Reality is objectively given as facts and can be measured as a set of properties. Different researchers will get the exact same results given the same techniques. Such studies often set out to test an hypothesis or a theory.

2. Interpretive paradigm: Reality can only be accessed through social constructions and the sharing of ideas and experiences. In the area of information systems, these studies generally attempt to understand how a system fits into a certain context.

3. Critical paradigm: Emancipation through social critique and trying to help eliminate the social problems encountered is a big part of critical studies. Within the field of information systems, participatory design is an often used technique where a problem is detailed and solved by a community.

4.2 Case study

A case study facilitates the research and exploration of an area within its context (Baxter & Jack, 2008) . Phenomena can be researched and dissected with regard to their social and cultural context. A case study is especially appropriate when studying the "how" and the "why" of a situation, or when believing the contextual conditions are of importance. In the area of information systems, the case study is the most commonly used qualitative methodology (M. D. Myers, 1997). Stake (2005) talk about three different kinds of case study research:

1. Intrinsic case study: The case itself is what's of importance, and there is done no attempt to generalize the information learned

2. Instrumental case study: The case itself is not as important as the broader issue that the case is used to represent and discuss

3. Collective case study: Differences and similarities between cases are explored, trying to replicate the findings across several cases

Which of these three types a specific case study research project occupies will help define the data gathering methods as well as how the researchers will work towards analyzing the data. The research in this thesis is an instrumental case study with roots in the interpretive paradigm. It is not the case of Ruter specifically that is of interest, but Ruter is used as a means to explore how automatic context-aware systems are understood and perceived by users. It gives the participants of the study a case and a system to discuss and relate to. Hence, this research is an instrumental case study rather than an intrinsic one. The specific methods used in this case study are observation through user testing of a prototype, followed by interviews and a focus group. Using such qualitative methods provides the researcher with an understanding of *why* the participants do as they do. This information can then be used to explain greater phenomena and compare the case study to others like it (Flyvbjerg, 2006).

4.3 Prototyping

A prototype is an unfinished version or a representation of a system created to test some aspects of the system. The later in development a user test is conducted, the more expensive it will be to make changes to the system (Mantei & Teorey, 1988, p. 438). Because of this one would prefer to be made aware of as many usability problems as possible early on in the development cycle. Even testing out the concept of the system before any code is written can be of great value. In the field of usability engineering, a prototype is referred to as a simulation of a systems layout and/or functionality, while for programmers a prototype can be a test program (Houde & Hill, 1997). Low fidelity (henceforth referred to as lo-fi) prototypes are those who look and behave less as a finished product, as opposed to high fidelity (henceforth referred to as hi-fi) prototypes, which are more complete and have a more realistic look to them and are more interactive. A foam figure or mockups created with pen and paper are considered prototypes just as well as a website being developed in HTML, CSS and Javascript, however they do have different applications and usages. Lo-fi prototypes generally take substantially less time to develop. While a hi-fi prototype can take weeks, if not more, to complete, a lo-fi prototype can be ready in less than a day. Additionally, a single bug in the hi-fi prototype can ruin the entire test, wasting resources (Rettig, 1994). For instance, Liu and Khooshabeh (2003) developed a lo-fi and a hi-fi prototype of the same system, where the former took one day to implement and the latter took two weeks.

> *To design well, designers must be willing to use different tools for different prototyping tasks. (Houde & Hill, 1997)*

When testing hi-fi prototypes, the users tend to think that the design closely represents the design of the finished product (Houde & Hill, 1997), so they might concentrate on issues related to details and nitpick on things like color, specific placements of objects et cetera. With lo-fi prototypes, which are less polished, users tend to focus on issues related to the concept, general layout, functionality and navigation (Rettig, 1994). As a result, lo-fi prototypes can be used effectively in the first parts of the development cycle, or when considering new functionality or other bigger additions to the system. The cost is fairly low compared to developing a hi-fi prototype. There are no clear distinguishing factors that decides if a prototype is lo-fi or hi-fi, it is more of a continuous scale than a binary one, but Houde & Hill defines fidelity as "closeness to the eventual design".

For ubiquitous systems specifically, Liu and Khooshabeh (2003) found that increasing the fidelity and the automation would result in "better usability data".

> *Paper screens were not of sufficiently high fidelity to convince participants that they could really incorporate Kitchen-Net in their lives. (Liu & Khooshabeh, 2003)*

The lo-fi paper prototype they created needed to be controlled by facilitators during testing, affecting the flow of interaction and reducing immersion. After having created a prototype, it is usually subjected to user testing.

4.4 User testing

In this thesis, the concept of testing a system on users will be referred to as "user testing", even though its often called "usability testing", as the two terms are often used interchangeably. User testing can be seen as a sort of observation technique, and can be executed in different ways, ranging from testing a complete or incomplete system with five users, to testing and getting feedback on a paper prototype with tens of users. Presenting the users with different tasks to complete through use of the system, while expressing their thoughts out loud, has long been used as one of the main methods of user testing. This helps the designers and developers see in what way the system matches the users way of thinking, their usage patterns and mental models (Kaikkonen, Kekäläinen, Cankar, Kallio, & Kankainen, 2005). It is nearly impossible to always predict what the users are going to do when faced with a task (Toftøy-Andersen & Wold, 2011).

> *The most effective way of understanding what works and what doesn't in an interface is to watch people use it. (Group, 2014)*

Testing with current or potential future users is a whole different task than developers and designers discussing, testing and designing it at the office. Developers and designers have been working on the system for some time, hence they have a deeper understanding of how its supposed to work than the layman or average user. Additionally, the designers have designed the system themselves, rendering them biased. The conceptual model of the system might be based on their mental models which may very well be completely different from the mental models of the users. Last but not least, people working in the development of IT systems are generally very "tech-savvy", as well as used and accustomed to the abstract thinking one needs to efficiently operate a computer. Many systems are used by both people who are comfortable with technology and people who are not, so programmers or designers often won't represent an average user.

The more clearly you have established the purpose of the user test, the easier it is to plan the test and interpret, review and use the results (Toftøy-Andersen & Wold, 2011, p. 39). When the purpose of the test has been decided, one can choose which functionality and parts of the system to be included in the test, to be able to narrow it down and keep it relevant. What question or hypothesis you want answered directly affects how your test should be organized and what kind of tasks you should give the testers.

When testing a system the users need to be given things to do - tasks to complete. The wording of the task should not reveal exactly how the task should be completed or tell the user to press this or that button. The task should be vague enough so that the users themselves can try to figure out exactly how to solve the task. In order to motivate the user, the task can be wrapped up in a scenario which offers a context and some meaning to the user, hopefully immersing them in the task. When they are performing a task it is beneficial that the user feels like this is something they could realistically do in their everyday lives. A task need not be longer than a sentence or two (Group, 2014).

4.4.1 Pilot testing

Having the usability specialist or a team member run through the test as he or she would be a tester will greatly help in reducing misunderstandings, poor choice of tasks and poor timing constraints when finally testing with users (Schade, 2015). Running a pilot test will help uncover problems with the testing plan that would otherwise have gone unnoticed while planning. Even a single pilot test can help increase the quality of future tests, especially when the person planning and conducting the test is new to the subject area or the field of usability in general (Rettig, 1994). Is the flow of tasks natural? Are the questions we plan on asking the right ones? Are the questions leading? Do we need a different audio recorder? Do we need more time for each task? Are we introducing the user to the test and application in a good way, helping them understand the context of and reason behind the test? These are the types of questions one can answer through one or two pilot tests. The ultimate goal of the pilot test is to improve the test procedure to allow it to run more smoothly when testing with real users so the data collected is relevant and usable (Toftøy-Andersen & Wold, 2011, p. 66).

4.4.2 User testing in context

According to Kaikkonen et al. (2005), when testing context-aware systems the result of the test is affected by where and in which context the tests are conducted. Most systems can be either tested in a lab, in their true realistic context, often called a field study or in-situ study, or somewhere in-between. However, the importance of context in testing grows for mobile systems, ubiquitous systems, or other highly context-sensitive systems such as medical emergency room equipment. Kjeldskov and Stage (2004, p. 600) talk about the three major issues related to testing in context:

1. "It can be complicated to establish realistic studies that capture key situations in the use-context [...]."
2. "It is far from trivial to apply established evaluation techniques such as observation and think-aloud when an evaluation is conducted in a field setting."
3. "Field evaluations complicate data collection and limits control since users are moving physically in an environment with a number of unknown variables potentially affecting the set-up."

When testing in a lab these problems are often more or less solved, however, at the cost of losing the realism. For stationary or less context-sensitive systems the researchers can introduce some realism to the lab by decorating it and possibly include sound clips and dress up. For mobile and other context-sensitive systems, however, the realism of and interaction with the physical environment as well as the movement of the user is more or less lost in the lab. Abowd and Mynatt (2000) argue that ubiquitous systems require testing and evaluation in their expected use context:

> *Deeper evaluation results require real use of a system, and this, in turn, requires a deployment into an authentic setting. The scaling dimensions that characterize ubicomp systems - device, space, people, or time - make it impossible to use traditional, contained usability laboratories. (Abowd & Mynatt, 2000, p. 51)*

Kjeldskov and Stage (2004) saw that participants testing a system while sitting still found more usability problems than those that tested the system while walking on a treadmill or down a pedestrian street. However, the problems found while sitting as opposed to walking were mostly cosmetic. Additionally, the workload that the participants perceived increased the more their movement during tasks increased.

In their paper "Is it worth the hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field", Kjeldskov, Skov, Als, and Höegh (2004) found that conducting a user test in the field often will consume more time than a test in the lab. Additionally, there were found significantly more usability problems in the lab than in the field tests. The only problem that was found in the field tests that was not found in the lab was one of data validity, indicating that the field tests allowed for a more realistic test setting. As they found more usability problems in the lab tests, and they were substantially less time consuming to conduct, they proposed not using expensive field studies to evaluate a system's usability. However, they did suggest that field studies could be the best option in early stages of design:

> *[...]field studies may instead be more suitable for obtaining insight needed to design the system right in the first place. (Kjeldskov et al., 2004, p. 70)*

Kjeldskov later highlighted this point when reviewing the debate his statements had sparked, saying that field studies indeed have their own applications, but won't necessarily substitute lab tests entirely (Kjeldskov & Skov, 2014).

In response to the study of Kjeldskov et al. (2004), Nielsen, Overgaard, Pedersen, Stage, and Stenild (2006) saw that significantly more usability problems were found in the field tests than in the lab tests, directly contradicting what Kjeldskov et al. (2004) had found. The field tests produced more issues relating to cognitive load and style of interaction, which may have been the cause of the more realistic context that the field study provided. As opposed to Kjeldskov et al. (2004), they concluded that field tests in context indeed is "worth the hassle".

Following on this response, Rogers et al. (2007) argue that field tests are especially important in the area of ubiquitous computing as it consists new and untouched interactions and experiences. They state that field tests give more answers to how users interact with, understand and experience the systems, which are important aspects in ubiquitous computing. This comes in contrast to studying the intricacies of the interface itself which often can be sufficiently done in a lab.

Kaikkonen et al. (2005) found that when testing the user interface and navigation scheme of an application, there are no major differences between a lab test and a field test. When looking at the general concept of the system, however, the users had an easier time expressing themselves and talked more freely after the field tests. Some social aspects such as checking out one's surroundings or looking at other passengers while waiting for the application, was participant behavior found in field tests but not in lab tests.

Kjeldskov and Skov (2014) argue that the main challenge of mobile usability engineering has evolved from studying the intricacies of the interface itself and developing on a small screen, to the act of integrating mobile systems into an ecosystem of different technologies and contexts, and how to design good interactive experiences. In such cases, they argue, a controlled and closed test lab might not fully facilitate the realism and context of the ecosystem the device is thought to operate in. In conclusion they mention that discussion of *when* and *how* field studies should be conducted is more appropriate than discussing *if* they should be conducted at all.

4.5 Interviews

An interview can be conducted in several different ways. In a structured interview the researcher reads questions off of a list, much like a questionnaire, and sticks strictly to this list of predefined questions. In a semi-structured interview the researcher has prepared a topic for discussion and some open-ended questions which might or might not be asked, and the order in which the questions are asked can be changed mid-interview. Additionally, in a semi-structured interview new questions might be formulated underway by the researcher based on the response of the participant. Lastly, an unstructured interview is where the researcher has not prepared any questions and often just a very general and broad area of discussion, is any at all (Crang & Cook, 2007, pp. 60-89). The more aware you are of the questions you want answers to, the easier it is to set the scope of the interview and find out which questions to ask. Interviews can be conducted in a way that lets the researcher purely collect data through asking questions and logging answers, barely - if at all - commenting on the answers or opening for a discussion. Some interviews, however, are conducted in a way where the researcher and the interviewee help each other construct an understanding. If the researcher wants to open up to more of a discussion and follow-up questions to really understand the interviewees viewpoint, he still needs to be careful not to ask any leading questions. Leading questions are questions where the way it is asked or formulated hints towards the answer the researcher is looking for, making the interviewee likely to answer accordingly.

Interviews can be recorded through note-taking, or on audio or video tape. While recording through voice one should still take notes of timestamp and inaudible signs such as body language, pointing, miming or demonstrations, as well as the researchers thoughts throughout the interview. Note taking also applies to video recorded interviews, helping the researcher more easily locate specific moments of interest of the interview or again writing down his or her own thoughts. Note taking can also help capture things that a non-perfect audio or video recorder can not, such as movements or body language outside of the line of sight of the camera, or sounds and distractions that the microphone won't pick up. Watching or listening to a recording might have the researcher mistakenly interpret a reaction to an outside interruption, such as a sound, as a response to the question. Looking at one's notes while watching or listening to a recording will help clear up mistakes like these. Too much writing of notes, however, can be very distracting in itself and might distract the interviewee from what he/she was trying to say, or might make the researcher miss something the interviewee was saying as his/her concentration was on the note-taking (Crang & Cook, 2007, p. 81).

4.6 Focus group

Allowing people to discuss, help and challenge each other through the social context of groups might produce new and different data than gotten from one-on-one interviews. The participants can discuss and share their experiences and react to each other's experiences - they focus on each other instead of the researcher. This in turn might uncover issues, praise or points that might not have been raised otherwise.

> *[...]The group should be used to encourage people to engage with one another, verbally formulate their ideas and draw out the cognitive structures which previously have been unarticulated. (Kitzinger, 1994, p. 106)*

The researcher or moderator should not ask each participant individually them same question one after the other, but can ask a question to the group as an entity and allow them to react to each other. As free flow of communication and an open area for addressing concerns or raising points is important, conducting a focus group with an already established group might prove itself useful. The size of the group can range from as small as 2 to 3 users to as much as ten to twelve. Too large groups, however, might reduce the time each participant has to and can speak, hence not giving them the opportunity to fully express their thoughts and opinions. On the other hand, in a small group the discussion may fade out more quickly and the researcher has to encourage the participants to keep talking. In such a case the researchers can then end up with little data. When the group gets too large it can happen that the participants automatically split themselves into sub-groups where the discussion is flowing only within the bounds of these sub-groups, effectively causing it to become two or more groups instead of one large (Crang & Cook, 2007, p. 94). From personal experience facilitating workshops, when the group gets bigger than six participants some of the more timid participants do not feel comfortable talking and hope they will go unnoticed if they do not say anything. This effect is strengthened if one or more of the other participants are especially vocal, confident, express authority or are aggressive. As Crang and Cook (2007, p. 96) states it: if a leading figure sets out an opinion early on it might establish a precedent for the rest of the focus group, and others "may essentially opt for the quiet life rather than face a possibly hostile reception to their views". This concept is known as a groupthink. If such a case would arise, the moderator can try to break up this groupthink by introducing a different viewpoint on the subject.

Using participants who are part of the same group and already know each other to different degrees can enable the participants to relate to each other and have a better flow in the discussion, as they feel more comfortable with each other. Additionally, they can contest or relate statements other participants have said to real life experiences that not just themselves, but also the participant uttering the statement have had. As they know each other they can challenge one another when they say they behave a specific way or believe a specific thing, but in reality behave in an entirely different way (Kitzinger, 1994). In general groups can sometimes actively help further discussion and allow timid or less-spoken people to speak their minds, as more outspoken and vocal participants "break the ice" or they disagree with someone else and feel the need to interject. This can be especially true for private and sensitive information on taboo topics as participants can feel like others support their views and experiences.

4.7 Reviewing and analyzing data

After having transcribed the user tests, interviews and focus group, more thorough work on the research documents is possible. Khandkar (2009) explains one method of doing this called open coding, which helps organize the statements and ideas of the interviews, as well as avoid imposing the researchers predefined and biased set of categories and thoughts on the research data. Reading the transcriptions thoroughly, preferably accompanied by listening to the audio tapes, will help the researcher remember the context of each conversation, idea and point. Especially the re-listening to recordings will help the researcher not get a too over-rationalized look at the data or forget about the fact that an interview is a cooperative intersubjective task, where ideas often surface through collaboration. The meaning and semantics of each statement is what's especially important to note in this stage of the review, both the interviewer and the interviewees statements as, again, interviews are an intersubjective task. According to Khandkar (2009), the idea at this point is to find the meaning of the specific statements, not the general discussed theme or connections between different statements, and then make a note of it for instance in the margin next to each statement. Then, after having completed this task, the researcher reviews these notes, and labels or "codes" the statements and parts of the notes that share similar semantics. These codes can be anything the researcher prefers, from color to lines to text. The parts can be as big or small as the researcher prefers, may vary from one word to a chunk of text throughout the review, and each part of text might be associated with several codes. Crang and Cook (2007, p. 139) argue that information such as in which context a statement was expressed, the inflection of the statement and any special body language might also be worth including in the codes from time to time.

When working through the material, the researcher should write each new code on a second piece of paper and explain what the code means and why it was created. This will help him remember what each code means when he forgets, as well as force him to think through each code and only create a new one if it is indeed needed, or maybe expand it if it was too small. Connections, ideas and broader themes, new or old, will be brought to light when coding, and these should be noted on yet another different piece of paper, which the researcher can and should look back at regularly. According to Crang and Cook (2007, p. 139), after all the notes have been coded it is time for a review of the codes themselves. Rewriting all the codes on a separate piece of paper helps the researcher organize them, double checking that no code has been written in two similar yet different ways so it appears to be two separate codes, and last but not least, find codes that should be combined into one because of their striking similarities and connections. To finish, Crang and Cook (2007, pp. 131-149) say that the codes and their connections must be reviewed, for instance by explicitly noting each code's occurrence and location through the materials on special index cards. This lets the researcher take a step back and look at how the statements, ideas, codes and chunks of text relate.

4.8 My use of the methods

My research started with the development of a prototype in the form of an Android application, which scanned for a specific set of Bluetooth beacons, and giving a notification that a ticket was purchased when the user came close enough to a beacon. That is, as long as the user had the application running, the ticket would be automatically purchased when approaching a beacon. The development process of this prototype is further detailed in Chapter 5. Following the development of the prototype, I tested it with 7 participants, 3 of which tested it in a lab, and 4 tested it in-situ in a field test on the metro in Oslo. Each of the seven participants completed two short tasks with the use of the developed application prototype and answered questions and expressed their opinions and impressions during a semi-structured interview. The first task was to walk to the metro station, a real metro station in the field test and represented by a metro sign in the lab test, and then view the ticket when the application had purchased one. That is, the participants simply walked to the station with the application turned on, and the ticket was automatically purchased. For the lab testers, the second task consisted of walking past the "metro station" and canceling the ticket the application would purchase, showing them that such functionality was indeed possible to implement. The field testers completed the second task by walking across a bridge over a metro station and then canceling the purchased ticket. On average a lab test and subsequent interview took 36 minutes, while an average field test with subsequent interview took 44 minutes. Figure 4.1 shows a picture from one of the lab tests on the left and one of the field tests on the right. Three out of the seven participants were computer science students (one Bachelor's student and two Master's students), one was a nursing student, one was an mechanical engineering student, and the remaining two were political science students.

Figure 4.1: Picture of one of the lab tests (left) and one of the field tests (right). In the lab test we can see the paper metro sign created after feedback from the pilot tests.

After these user tests, a focus group was held with 4 participants discussing automation of tasks with the use of QR codes, NFC and Bluetooth, as well as the participants' experiences with Bluetooth. These participants did not take part of the previous user tests. In order to avoid groupthink, the participants wrote down their immediate thoughts about QR codes, NFC and Bluetooth on paper, and their thoughts, impressions and experiences were discussed afterwards. The focus group lasted for 50 minutes. During the focus group, two participants were computer science students, one was a teaching student, and the last was a marine engineering student.

Spiekermann (2008) mentions that prototypes of ubiquitous systems need to be fairly advanced before they can be tested with users, enabling a proper portrayal of the concept, new technology and mode of interaction. As a result, developing and deploying prototypes can be difficult and time consuming, and Spiekermann argues that prospective findings can be just as important as retrospective findings. After all, with new and innovative technologies such as those in ubiquitous computing, what Spiekermann calls the users' "pre-purchase intentions" towards a system are of great importance. A negative attitude towards the concept will result in people not purchasing the system. This thesis discusses both prospective and retrospective findings and attitudes, as some participants had a hands-on experience with the prototype and some did not. Additionally, some findings, such as time it would take to fully trust the application, were of a prospective nature for both the focus group and user test participants.

The user tests, interviews and the focus group were all recorded on an audio device, and the recordings were transcribed. After having transcribed the recordings I read through them with the audio tape running, underlining and writing notes in the margin when I encountered parts of special significance. Using open coding to analyze the gathered data, I read through the transcriptions again without the audio tape running, writing down labels that I felt would help me organize the data. Examples of these labels were (translated from Norwegian) "Power-fear", "Power OK", "BT no point", and "Trust Ruter". These labels were written both at the front of each transcription and wherever in the text they were relevant, sometimes with explanations for myself as to why I found it interesting. This helped me organize the data and easily see which participant felt and had said what, and how many participants shared the same views. Before I could record their voices on audio, all participants signed a paper of informed consent, agreeing to the audio recording, given that I would delete them after the research was concluded.

The participants of the user tests were selected by the use of the snowball sampling method, where participants put the researcher in contact with subsequent participants. As snowballing can lead to many participants with similar traits and interests, the participants were explicitly told not to think too much through who to suggest, and that the participants did not have to be particularly good with computers. The participants of the focus group were a group of friends known to the researcher. This was to allow the focus group to run more smoothly and allowing participants to remind each other of incidents, call each other out when they said something wrong about their own behavior et cetera. All in all there were 12 participants in this study. Two limitations should be brought forward: (1) of the 12 participants there was only 1 female, and (2) the age of the participants were between 22 and 25 years. During further research the demographic should be more diverse. After all, Spiekermann (2008, p. 110) noticed that people below the age of 30 generally perceived less control over the tested systems than those older than 50, so examining a larger variety of users can be of interest. As for the number of participants in my study, I started to notice less new data during the sixth interview, and concluded by the seventh that I had reached data saturation. That is, I was satisfied with the data gathered and I concluded that further tests and interviews would be of less value. As a result, the focus group discussed some topics the user tests and interviews had not, generating some new data. However, the data from the focus group was very much in line with what had been found during the user tests and interviews, although more towards a negative attitude regarding the system.

Chapter 5

Prototype

This chapter details the process of developing the prototype, provides info on the beacons used as well as functionality and screenshots of the prototype.

5.1 Development process

The prototype, an Android smartphone application, was developed over the course of eight weeks, although with my attention split between writing on the thesis and developing the prototype. The application was built to run on Android phones with Bluetooth 4.0 capabilities, and the app's minimum Software Development Kit (SDK) requirement is version 21 (Android 5.0 "Lollipop") and target SDK version 23 (Android 6.0 "Marshmallow"). Figure 5.1 shows a rough activity diagram of the application.

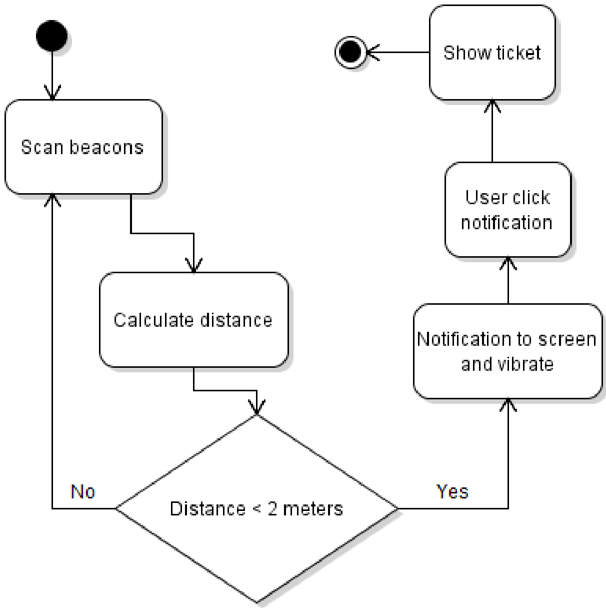

Figure 5.1: Activity diagram of the developed prototype

The distance calculation is executed with a formula collected from the open source Android Beacon Library[1] created by Radius Networks[2] and published under the Apache 2.0 License. This formula calculates the distance from the beacon to the phone with the use of the RSSI (Received Signal Strength Indicator) and a value configured into the beacon itself saying what the RSSI is known to be at a distance of 1 meters, called the txPower. After having trouble configuring the txPower values onto the beacons, I hard coded these values in the application, which was adequate for the purpose of this thesis. As the RSSI values from the beacon readings tended to fluctuate, I decided to combine several readings in order to get a more accurate distance estimation, as suggested by B. Cook et al. (2005). At first I implemented a low-pass filter in order to "smooth out" the readings and reduce the impact of large fluctuations. This gave a good distance estimation, however, it was not very responsive. That is, the time it took to calculate the distance from a beacon was deemed too long, as it was about 12 seconds. This comes from

[1] www.github.com/AltBeacon/android-beacon-library/blob/master/src/main/java/org/altbeacon/beacon/distance/CurveFittedDistanceCalculator.java

[2] www.radiusnetworks.com

the fact that beacons broadcast at a maximum rate of 100 ms, not giving the low-pass filter enough input values. I ended up taking the advice of B. Cook et al. (2005), averaging the RSSI value over 5 different readings. Before including a reading in this average I checked to see if its RSSI value deviated from the current average reading by more than 10 dBm. If it did, I discarded it and waited for the next reading. This was done in order to further combat fluctuations, however, at the cost of responsiveness. The beacons seemed to broadcast packets in bursts and at irregular intervals, even though I updated their firmware to tell them to broadcast a packet every 100 ms. As a result, the speed of distance estimation and the application's ability to react to change was not as quick as I had hoped, however, it was adequate for the purpose of this thesis.

The application scanned for the beacons based on static beacon address, meaning that without access to those particular beacons the prototype is of little value.

5.2 Texas Instruments SensorTag CC2541

The beacons I used were SensorTag CC2541's[3] from Texas Instruments. The device's low-power mode and short mode-transition time causes it to consume little power. Several different sensors allow the CC2541 to measure environmental conditions such as heat, humidity, and acceleration, however only the broadcasting of device ID and name was actively used by the prototype in this thesis. The CC2541 uses a single 3 volt CR2032 button cell battery.

[3] www.ti.com/tool/cc2541dk-sensor

Figure 5.2: One of the TI SensorTag CC2541 Bluetooth beacons used in this thesis

By default the CC2541 discontinued its broadcasting procedure after 60 seconds, only continuing if a button on the side of the beacon was pressed. In order to broadcast indefinitely until either the start/stop button was pressed again or the battery ran out, I updated the firmware.

5.3 Functionality

The app consists of two Android activities[4], or screens, one to start or stop the BLE scan and the other to view one's expired and unexpired tickets. The first activity consists of a start button, a stop button, a button to access the tickets activity, and a text area showing whether or not the app is currently scanning for BLE devices. Screenshots of these activities can be seen in figure 5.3.

[4]developer.android.com/guide/components/activities.html

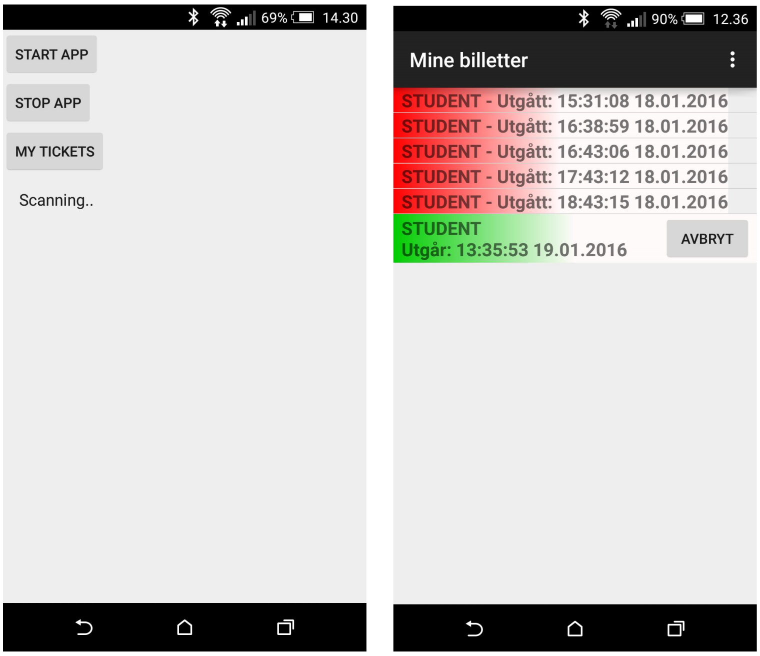

Figure 5.3: The main Android activity displayed upon startup (left) and the My Tickets activity (right)

When the start button is pressed, the BLE scan starts with a callback function, which is called each time a device is scanned. This callback function checks if the scanned Bluetooth device is one of the SensorTag beacons and if it is, calculates the distance between the phone and the beacon. If the beacon is estimated to be within 2 meters of the phone, a vibrating high priority notification is displayed on screen, letting the user know that a ticket has been purchased and that the notification can be clicked to view further details. Clicking this notification takes the user to the MyTickets activity, displaying the newly purchased ticket as well as previously purchased tickets, regardless of them being expired or not. If a ticket has been purchased within the last minute, a button to cancel the purchase is displayed next to it. This button is automatically removed after the minute has passed even if the user is active on the MyTickets activity at that time. The decision to remove the button instead of disabling it was made in order to avoid clutter and to clearly show that once a ticket is no longer cancelable it can never return to a state where it can be canceled. When a ticket is canceled the entire scanning process is paused for 60 seconds, displaying this information in a message on screen. When a ticket is purchased, the BLE scan is automatically stopped so as not to purchase any more tickets or consume any more power. After ticket expiration the scan is resumed.

The application's GUI is not polished, and contains few colors and details in order to keep the users' focus on the concept and functionality of the application, not the details and intricacies of the graphical design, as explained in 4.3.

Chapter 6

Findings

This chapter details the findings from applying the methods, i.e. the user test, interviews and focus group, to the case using the prototype. First, the results and effects of the pilot tests are briefly mentioned. Second, the findings from the user tests and accompanying interviews are detailed. Lastly, the findings from the focus group are presented.

6.1 Pilot tests

Before the application was properly tested with users, it was tested on two fellow interaction design Master's students individually, in order to get feedback on and discuss the execution and tasks of the user test. This was to ensure the test was running smoothly, and that the data gathered was relevant and usable.

The first pilot test showed that the application was not polished, and that parts of the in-development debugging tools were not removed. This confused the tester and he wondered why the information was pointless to him. A notification with the text "Ruter-billett: 90:59:AF:0B:8C:17 - Distance: 1.654754m" created more confusion. Additionally, for debugging purposes, the app logged each beacon scan, filling up the screen when the tester had the app open, which resulted in more confusion as to if this information was relevant to him. The introduction of the application as well as the goals and purpose of the test was not clear and thorough enough, leaving the tester worried as to exactly why he was there. He advised me to ask some introductory questions in order to help ease the mood and better establish the setting of the test. Lastly, the flow of the tasks did not run as smoothly as desired. The first task did not have a clear end condition and was nothing more than for the tester to go to the area representing a station and wait for a ticket. After the ticket was purchased the task was completed, but as the user didn't interact with the app through any clicks, it was unclear to him whether the task was completed or not. The second and last task was more or less the same task again, but the tester was asked to check if his ticket had run out or not before going to the "metro station". Again, the task had no definite and clear end, making it flow poorly. Between the two pilot tests this task was changed. The new version of this task told the participant to cancel a ticket that was purchased within the last minute. In order to let the tasks flow better in general, and make it easier to understand where the metro station was, a copy of a metro station sign was created out of cardboard in a mix of colors. This sign can be seen in figure 4.1 on page 60.

The second pilot test ran more smoothly, and as the second pilot tester had an Android phone the app was successfully tested on a different phone than the one used in development. The metro station sign allowed the tester to find the metro station without the researcher having to lead him to it, the first task had a more clear finish condition as the tester was urged to look at his ticket after purchase, and the last task was changed completely. As a result of this second pilot test I rearranged the planned order the questions of the semi-structured interview, prepared a better introductory statement to read to participants, and prepared something to say between the two tasks so the flow of the test seemed more natural.

The pilot tests helped me prepare the introduction, task and questions for conducting the "real" tests. A significant improvement of flow, as well as proper introduction to and explanation of the project to participants was achieved.

6.2 User tests and interviews

The first three user tests were carried out in a large classroom at the University of Oslo, and the subsequent four user tests were carried out in-situ on the metro in Oslo. Each of the seven participants completed two short tasks with the use of the developed application prototype and answered questions and expressed their opinions and impressions during a semi-structured interview. The first task was to walk to the metro station, a real metro station in the field test and represented by a metro sign in the lab test, and then view the ticket when the application had purchased one. For the lab testers, the second task consisted of walking past the "metro station" and canceling the ticket the application would purchase, showing them that this would indeed be possible. The field testers completed the second task by walking across a bridge over a metro station and then canceling the ticket purchased.

6.2.1 Experiences with Bluetooth

Every participant had some previous experience with Bluetooth and had some predetermined associations. However, only three out of the seven participants had heard about beacons before. Of these three, one had only heard the name, one was working with beacons for his masters thesis, and the third had heard that it could be used by advertising agencies. All the participants had used Bluetooth in order to connect their phone to either speakers or headphones to listen to music, one on a daily basis, some on a monthly basis and some only a couple of times throughout their lives. Two of the participants used Bluetooth to connect their phone handsfree to the car on a regular basis, and one of these had Bluetooth constantly on in order to connect his phone to his car each time he entered it. Lastly, four of the participants had used Bluetooth to transfer files such as images and music between their phone and a friends phone before smartphones and pervasive Internet access emerged.

One participant felt that the use of Bluetooth and wireless technologies in general was very cumbersome, especially in the context of home technologies, and consistently got a better experience when using wired technologies where possible. Examples were wired compared to wireless printers, wired compared to wireless networks and television signals over wire such as an HDMI cable compared to wirelessly through a Chromecast. He stated:

> *I'm not that big a fan of [wireless technologies] in general, as I've had some bad experiences with it. Sometimes it works, sometimes it does not work, and then it just becomes annoying.*

He and one other participant described themselves as inherently skeptical to new applications and technologies even though they also described themselves as fairly interested in technology. Often they would look at new technologies as fads and simply the result of good marketing. Over time, however, seeing that a technology was not a fad after all, they could then decide to give it a try. Both mentioned QR codes as one of the technologies they after some time tried to use a few times, but they soon stopped as they saw little gain in using it. An application they both mentioned they firstly saw as just a fad was Vipps[1] - an application for transferring money from your bank account to a friends back account only by name or phone number. But they ultimately began using extensively because of its usefulness. One of these participants talked about the state of clutter and mess his phone would be in if he were to download every new trending application, and felt that he most times only used them a couple of times anyway. They had both had poor experiences with their Bluetooth devices not working consistently.

This shows that they all had some experience with Bluetooth, although mostly of limited extent. They had mostly used Bluetooth to connect to speakers, but only a few of them at a regular or semi-regular basis.

6.2.2 Ticket on travelcard or phone

Different participants usually purchased tickets for public transportation through different channels. They all used single tickets on travelcards only if there was no other option, e.g. if they were afraid their phone's battery would not last the entire journey. Two of them had 30-day ticket on their

[1] www.vipps.no

phones, two had 30-day ticket on travelcards, and the remaining three purchased single tickets on their phones. They all mentioned running out of battery as a drawback of having their ticket on their phone, but for single tickets they usually knew whether or not their phone would last for the hour needed. Two participants, one with his 30-day ticket on a travelcard and one with the 30-day ticket on his phone, said that the main problem of the phone was that it needed to be free of errors for a whole month. If not they risk losing their ticket. The same two persons said that having the ticket on the phone was incredibly convenient and that it was way easier to purchase a ticket in an app. However, neither of them felt they could trust the phone for an entire month. Nevertheless, one of them had his 30-day ticket on his phone, even though he explicitly said he doesn't like it. The other stated that he would love to be able to trust his phone for an entire month:

> *"It would have been very lovely if I trusted it [for an entire month]. So I might do that eventually."*

They all said that they were always carrying their phones, and this was the main advantage of the phone compared to a travelcard, both for 30-day and single tickets. Additionally, it was mentioned as a more convenient way to purchase the ticket, instead of having to walk up to a machine which may be out of order. The person who had a 30-day ticket on his phone, yet did not like it, said that convenience was the reason he used the phone.

> *"I don't like it, I actually strongly dislike it. [...]If it had been just as easy to purchase a ticket on the travelcard I would have. I prefer that, because I've run out of power on my phone before, and then you're in trouble"*

He had tried a single ticket on his phone once, downloading the app to do it, and from there the threshold to move his 30-day ticket to his phone was substantially lowered.

6.2.3 Bluetooth power consumption

Five out of the seven participants were afraid that having Bluetooth enabled on their phone would consume a substantial amount of battery. None of these five had Bluetooth enabled unless they were actually using it. Of the two other participants, only one had Bluetooth enabled at all times because he used it on a daily basis, and the other participant decided to keep his Bluetooth disabled because he only used it a few times a week. The latter said:

> *"It's my old way of thinking. I try telling myself all the time '[Bluetooth] doesn't consume any power', and then I turn it off."*

When asked why they did not have Bluetooth always enabled, even though some of them used it on a semi-regular basis, the general response was a combination of "Why would I ?" and "Im afraid it consumes too much power". Many of them were dumbfounded when asked why they don't have Bluetooth enabled, like it was a question where the only obvious answer was "Why would I?". One participant said:

> *"What is really the point? Let's say that throughout the day, Bluetooth consumes 1 % power. Let's just say that. That 1 % might potentially be what makes the difference of your phone running out of battery or not. Then it's not worth it if you don't need Bluetooth at all. But if you used it several times a day, and it still only took 1 %, it would be worth it."*.

This particular participant used Bluetooth a few times a month when taking his car for a long drive or connecting his phone to his speakers, not seeing any reason to have Bluetooth enabled at other times. Additionally, two participants mentioned that the Bluetooth icon that appears in the phones top bar was taking up too much space and that it made the bar messy.

Several participants admitted that they were not really sure whether or not Bluetooth consumes a substantial amount of power, and that technology might have come a long way in the last few years. However, they all talked about remembering Bluetooth consuming a lot of power when they were younger. When I asked *why* they didn't have Bluetooth constantly enabled, one participant said:

> *"It consumes power. That's what I've heard. The media has planted that idea in my head".*

Another participant stated:

> *"It might be that I have an old fashioned interpretation of Bluetooth. It might be that it used to consume more power and memory then than it does now. Now it might be running more in the background. But it used to consume battery, that's the main reason I don't use it."*

A third participant said:

> *"I also feel like it consumes power, I don't know how much it consumes, but at least it used to, I don't know if it does any longer. I don't use it for anything, so there's no point having it enabled."*

Most participants explicitly stated that they might be wrong because they know technology is rapidly evolving, however, they still felt that a substantial amount of power was consumed by Bluetooth. Despite not having Bluetooth enabled mostly because of fear of power consumption, six out of the seven participants had their 3G/4G mobile connection enabled at all times, while the seventh participant only had it disabled because his phone's battery was damaged and barely worked. Their reasoning was that they used it extensively, constantly checking their phones, especially while using public transportation.

6.2.4 Reliability of the system

Most of the participants seemed surprised and impressed when the application vibrated and gave them a notification that a ticket was purchased. After the ticket was purchased, all four participants of the field test asked how the phone knew when we were on the metro station, three of them wondering if it was because it was using GPS coordinates:

> *"Is [the app] synchronised with Google Maps or something, so that when we come [to the station] it gets the GPS coordinates?"*

They asked this despite having been told in the introduction to the test that Bluetooth needed to be enabled, as Bluetooth was used to check when we were on the station. One participant asked if he had to find the beacons "like a Geocache"[2], which involve searching for packages with the help of GPS. Even after having explained twice to this particular participant that Bluetooth was used instead of GPS, he later asked where the beacon was located and again if it used GPS to locate him. He also struggled in grasping the concept of automation at the start, asking if he was supposed to order the ticket beforehand.

When asked if they would be able to trust that the application purchased a ticket each time they went to the metro station, the answers were split. Some were afraid the application would not correctly determine when they were on the station, and as a result not purchase a ticket:

> *"I would be dependent on the technology. I'd be dependent on [the application] doing it for me. And it is pretty easy to just press 'Purchase ticket'."*

Others were afraid the application would think they were located on the station even when they were not, and purchase a ticket just because they were somewhat close. Only one participant, who was the person who used Bluetooth the second most, had complete confidence in the technology. He fully believed that the only way the application would purchase a ticket when you did not intend to was by human fault when, for instance, crossing a bridge over a metro station like in the user test and the user did not notice the phone vibrating, so he could not cancel the ticket. He believed he might forget to cancel unwanted tickets from such scenarios from time to time:

[2] www.geocaching.com

> *"I trust 100 % in that [the application] is going to purchase the ticket. Theres no doubt about that. [. . .]I'm just afraid I'm gonna be forgetful."*

The ones who were afraid the application would purchase tickets when they were not needed, raised concerns about the task of canceling an unwanted ticket. Two said that the 60 second cancellation window used in the user tests would be too small, especially because they might not notice that the phone vibrated. The participant with the most Bluetooth experience and who had his Bluetooth on at all times was afraid that the application would purchase a ticket for instance if he drove over a metro tunnel or across a bridge over a bus. He suggested that instead of the application purchasing a ticket automatically, it should ask you if you want to purchase a ticket, and you should be able to click "OK".

> *"I'd rather have to press 'OK' instead of having to press 'Cancel'. 'OK' if I want a ticket."*

I mentioned this point to two participants in the subsequent user tests as they expressed concerns about too many unwanted tickets being purchased. One of them said it could be a good idea, but mentioned:

> *"Then you have to do more yourself [. . .]. A part of the point is gone. The point [of the application] is that you can forget about it and still have a ticket."*

The other of them thought this sounded like a good idea and that it should be like that by default, however, that one could customize the application to automatically purchase without asking:

> *"At least ask. Or that you can choose yourself, but as a standard setting it asks you [before it purchases]. Custom settings. Are we not in 2016?"*

This particular participant, who daily passed a metro station at a distance of about five meters,was afraid that the application would erroneously purchase a ticket for him each time. This would force him to either cancel a ticket or disable Bluetooth twice a day. Hence, he felt that having to confirm the purchase would fit his use case best.

Three participants mentioned that it would take some time for them to feel comfortable with fully leaving the purchase process in the hands of the application. It would take time for them to feel confident that the application would work consistently. One of them stated that he fully believed the technology would work, but he would double check just in case:

> *"I think, because I am one of those old-fashioned people, that I would have checked to make sure anyway, do you know what I mean? [...]Double check. Then after a week I would have thought 'Ah, it works'"*

A different participant stated:

> *"I would have been unsure at first and I would probably have walked closer to the beacons and so on. If I get a notification at that point, then it's ok, then you see that it has made a purchase"*

This shows that he most likely thought the beacons would be clearly visible to the public. A third participant said that the application needed to be entirely consistent for him to be able to rely on it:

> *"If it had worked all the time I think that I could have gradually lowered my shoulders and thought 'Yes, it works.' Its like a person thats just gotten a smartphone and is 60 years old and thinks 'I don't like this, this is difficult', but after a while they get a hang of it and it's going ok and they think 'This isn't so bad' and 'This is ok'. If it works every time I can trust it. But the second there is as much as one occurrence where it does not work, I will lose all my confidence in it".*

This shows that the application would have to be completely reliable in order for him to use it. This notion was strongly affected by the fact that money was involved.

6.2.5 Trusting the system

Five participants noticed that the symbol for Ruter, the company in charge of the metro, bus and tram systems in Oslo and Akershus, was displayed on the application and on the notification. When asked whether or not this would have any impact on their decision to use the application, all seven participants concluded after some thought that the application had to at least be supported by Ruter in some way. As a response to my question, a participant said:

> *"I would not have downloaded an application that didnt say 'Ruter'"*

Only one participant, however, said that he would actively research whether or not it was backed by them. Advertisement by Ruter on the metro, bus or tram, the Ruter icon and name on the application, and additional functionality of the pre-existing official Ruter app was mentioned by the other participants as ways to know that Ruter had backed the system. Being developed by Ruter or a different company was not important to any of the participants, as they all were aware that such companies often did not develop their own applications. What was important was for the application to be supported by Ruter. The same participant who said that consistency is key and that a single slip-up from the application would result in him losing all trust, said:

> *"If it's supported by Ruter it should work. Because I expect Ruter to have some sort of quality."*

The four participants that tested the application on the metro voiced concerns for legitimacy of the ticket, all four of them saying that if it was a part of a Ruter app they would know that it was legitimate, as opposed to if it was released solely by a third party. Being supported or released by Ruter, the application would make the participants more confident in that it would be reliable and work consistently. Two of them mentioned that if something still went wrong they would be able to get better customer support from Ruter than from a third party company.

> *"I feel that if Ruter screw up by giving me 60 tickets at the same time I can file a complaint. But I can't send an email to you, or some random strangers who made the application, and say 'God damn, you owe me 6000 kroner'."*

For all they knew, a third party could have stopped maintaining the application and had no interest in customer support if it failed. They meant that a failing application supported or developed by Ruter would have a smaller chance of getting them fined as the ticket inspectors would have already been informed if something was wrong with the system on that particular day. The participant who also wrote a thesis on the subject of Bluetooth beacons tested this application in the lab. He said:

> *"I can rely on it because it says 'Ruter' as well, and I know I can rely on Ruter. It gives it some credibility"*

One participant mentioned that using a third party app was undesirable because of the risk of identity theft. Two participants were afraid that a third party application could steal their money. None of the seven participants were not afraid that Ruter would steal their money, as they all had previously used the Ruter application to buy tickets:

> *"For me the fact that it's not developed by a third party matters. I trust Ruter, they don't do anything bad with my money. I assume."*

Lastly, two participants concluded that the functionality had to be part of the already existing official Ruter application for them to use, it in order to avoid cluttering up their phone with many different apps. They also said they know that the Ruter application already works, which would help their confidence in the new functionality.

> *"I prefer the official thing because then I know its the real deal. [...]If something is wrong, I don't know what it could have been, either that it doesn't purchase a ticket or charges me twice, then*

> *it can be fixed by Ruter. If the application was not developed by Ruter, the company could say 'Thats not our responsibility, contact [Ruter]'".*

6.2.6 Loss of control

The participants' opinions differed when asked whether or not the automation and accompanying loss of control would feel intimidating or liberating. Several participants felt that the autonomy of the application was more intimidating because money was involved. However, as mentioned earlier, if Ruter was in charge of the system they believed they could get their money back if something went wrong. The participant with the most Bluetooth experience excitedly said that it would be incredibly liberating, given that he would have to press an "OK"-button for the purchase to go through. He had earlier talked about the ticket system having gotten so convenient that he sometimes forgot to buy a ticket on his phone, and with this new type of system he would be reminded each time he needed a ticket. If the application did not prompt the user with an "OK"-button, he would have found it more intimidating than liberating:

> *"No, I wouldn't have felt safe using it. I wouldn't have managed to cancel the tickets in time. Maybe I didn't feel the phone vibrate. [...]I'd rather have to press 'OK' instead of having to press 'Cancel'. 'OK' if I want a ticket."*

A different participant, who often found the activity of purchasing a ticket cumbersome, would see this loss of control as fairly liberating and pleasant if he could be able to trust that his phone would be free of errors for a month. If not, he could not trust his 30-day ticket to a phone, and had to keep it on a physical travelcard. For single tickets, however, he thought this would be very liberating, as he himself on numerous occasions had either forgotten to purchase a ticket or the ticket machine has been out of order.

> *"Liberating. Because I forget it from time to time. You have to [scan] single tickets too, which messes with my head completely. I am intimidated by many other things, but not this."*

He was the one who seemed most ecstatic about the application, eager to use it and one of the participants who trusted the technology the most. There seemed to be no doubt in his mind that the application would only be liberating to him, not even remotely intimidating. But he still had to double check that it indeed did purchase a ticket, at least for the first week or two.

As previously mentioned, two other participants said that the loss of control would have been stressful, but after checking their phones each time they went on the metro for a few weeks, it might become more or less liberating, although still with a lingering skepticism. One of them mentioned that the vibration when a ticket was purchased would give him at least some sense of control because he would be notified whether or not it was working properly.

> *"I'd say that it's best that it vibrates, not even for such a long time. Maybe even make a sound. Because then you know that you actually have purchased a ticket. [...]Then you can walk with it in your pocket and feel it vibrate and think 'Ah, yes, thats probably a ticket purchase'. You don't have to bring the phone out."*

Then, as soon as something went wrong and it did not work properly, it would simply become a stress factor and the loss of control would become more intimidating than liberating. This was enhanced by the fact that the application dealt with their personal money.

The participant who was the most fond of Bluetooth technology in general, and who used Bluetooth the second most, found the application nothing but liberating. When told that some other participants were afraid that tickets would be purchased if they went too close to a station, and that they wondered what would happen if you were to only say goodbye to someone at the station, he replied:

> *"It's not intimidating at all. You never go to the station without taking the metro, anyway. And sometimes it's a hassle to walk up to the NFC reader to scan the card, and sometimes the card is old and you have to bend it so the chip comes close to the scanner. [...]Then this Bluetooth app is way better."*

Three participants admitted having been fined for not having bought a ticket, and they all said that using such an app would have helped them avoid the fine. One of them had on more than one occasion forgotten to purchase a ticket because he was distracted by his phone, and this once resulted in a fine. He was the participant who argued for the implementation of an "OK"-button, and said that it would at least remind him to purchase a ticket. He said:

> *"The worst part of having the ticket on the phone is that you completely forget about purchasing the ticket".*

A different participant said that he sometimes forgot to bring his travelcard from home, and got fined for it once. As his phone is with him at all times he would have avoided a fine with the use of the application. The last of the three participants had recently gotten a fine because he forgot to purchase a ticket in a state of intoxication. Again, they all said that an automatic ticket purchase application would have made them avoid a fine, as they carried their phones with them, just forgot to bring or purchase a ticket.

6.2.7 Security concerns

Three participants expressed concerns of someone breaking into their device through Bluetooth, as two of them had experienced exactly this about 10 years ago in middle school.

> *"Is it easy to hack Bluetooth now? Is it safe? People can abuse it through another person's phone, make people buy tickets that they do'nt want to have or something."*

Additionally, devices placed by third parties with malicious intent was a concern for two of them. Examples mentioned were jammers making the application unable to localize the metro station, duplicates of the metro beacons being placed somewhere else in town, or a device making your purchase numerous tickets hence wasting your money. This was only a major concern for one participant:

> *"Suddenly someone hacks Bluetooth, right. On the station. Or comes with an external device and 'blasts' my credit card."*

Of these three participants, one said it was not really a big concern of his, and that it was simply a thought that popped into his head. Another of the three mentioned that the Bluetooth security might have improved in recent years, and the third participant was very concerned as he was afraid of someone stealing either his identity or money.

6.3 Focus group

A focus group with four participants was conducted in a classroom. Three male participants and one female participant. The men were a computer science student, a building engineer student and a cybernetics student, and the woman was a teaching student. All the participants knew each other beforehand, which was a conscious choice in order to let the conversation flow easier, and for different participants to remind each other of and elaborate on each others experiences. The focus group lasted for an hour, being introduced with a discussion of the participants experiences and feelings on Bluetooth, NFC and QR codes, and talks about their similarities and differences, pros and cons. Next the topic of discussion moved to automation of everyday tasks, with a special emphasis on tasks involving money and transactions. In order to avoid groupthink, the participants took two minutes at the start of the focus group to write down their immediate thoughts about NFC, QR and Bluetooth.

The focus group contained both the participant throughout the entire research who was the least ecstatic about and was least knowledgeable about technology and automation, as well as the participant throughout the entire research who looked the most forward to a sort of utopian world where we could "live seamlessly". This resulted in a discussion with experiences, meanings and points from both sides of the spectrum.

6.3.1 Bluetooth, NFC and QR compared

The participants all had some experience with Bluetooth, however the one with the least experience had tried and failed to connect her phone to her Bluetooth earplugs, resulting in her seeing Bluetooth as too complicated. She had, however, used Bluetooth to take phone calls in someone else's car when others had initiated the connection, and found it usable:

> *"Mom uses Bluetooth to connect her phone to her car to receive phone calls! [...]I have done that a couple of times and its very good. I think thats really cool. [...]She just presses a button, I dont know how to connect it."*

The others had all used Bluetooth to transfer files between phones before the outbreak of pervasive WiFi, like many of the participants in the user tests, and one of them connected his phone to Bluetooth speakers a few times a year. None of them ever had Bluetooth enabled, except for when the one participant connected his phone to his Bluetooth speakers. Two of them felt that most of what you could accomplish with Bluetooth you could accomplish with WiFi:

> *"I'm thinking that you're always connected to the Internet, and there's many things you can do on the Internet that you cant do with Bluetooth. And then it just becomes a hassle to enable Bluetooth and connect the devices, instead of doing it over WiFi which is already enabled"*

Having to turn Bluetooth on every time they were to use it was seen as pointless when the same task could be accomplished over WiFi. When asked why they did not have Bluetooth constantly enabled, two participants mentioned both power consumption and the fear of becoming infertile as a cause of radiation, while the female participant said:

> *"I can't get it to work. I think it's a bit too complicated. For non-technological girls like me."*

They all eventually concluded that they did not use Bluetooth for any purposes, so they saw no point in having it enabled:

> *"It's really just that I don't have any need for Bluetooth. Why would I have it enabled if I don't use it?"*

Quickly followed by another participant:

> *"Yes, that's really it. I don't use it for anything"*

Additionally, concerns were voiced regarding the security of Bluetooth, two participants in particular fearing that whoever was in their proximity could hack into their phone as long as they had Bluetooth enabled. As for beacons, one participant knew that they could be used for advertisements, two of them had heard of them but knew nothing about it, and the last participant had never heard about Bluetooth beacons.

They had all used NFC extensively through bus passes, access cards to their university, their gym and other authentication purposes, one of the participants stating:

> *"I've had buss passes and some other cards like that, I actually like it. Its nice. It's nice to just walk up and scan it. It's quick and easy"*

Three of them had used NFC to make payments through a service called Cashless, an NFC card to which you have to transfer money before you can use it, and which is only usable in a few select shops. They had not used Cashless anywhere else than on concerts and festivals to purchase drinks and food. The fact that the Cashless card usually only contain a relatively small amount of money, as well as only being usable in a few select areas made the participants feel more confident in Cashless working properly and felt the risks were not as high as with an NFC debit card. The participant with the least technology experience and knowledge mentioned that it would be nice to be able to use NFC with your debit card like they do with the Cashless card, but she meant that would not be possible because of the lack of security:

> *"I considered saying earlier that NFC can be used to pay in the store, but that wouldn't work as you would need more security. It would have been incredibly nice to just put your card close to the terminal and then you've paid"*

Two other participants then quickly mentioned that this is indeed possible today, however, only on purchases below 200 Norwegian kroners. Three of them had an NFC chip on their debit cards, however, neither of them had ever used this chip to make a payment. They found the use of smart chip and PIN code to be sufficiently fast and easy, not having felt the need to try out the NFC chip.

> *"I'm so quick at writing my PIN code, so there's no point [in using NFC]"*

One of them mentioned that the 200 kroner limit was one of the reasons he had not taken advantage of the NFC chip in his debit card. All in all, NFC authentication was seen as convenient and easy, however security concerns and the fact that they did not feel like it offered all that much compared to the chip and PIN approach had resulted in them not using it for anything else than access cards and the Cashless card payment. One participant suggested that all credit cards and access cards could be stored on a single NFC chip card, and another participant interjected that such a functionality could have been integrated with a smartphone application:

> *"I have an idea. Forget about QR, forget about scanning. Give everyone a card, a scanning card, just like a credit card, only it's a card you can scan where you have bus passes, when you purchase a movie ticket it gets put on that card and so on"*

Neither of them had heard of Google Wallet[3] or Apple Pay[4] when I brought those up as similar applications to that idea. On the topic of QR codes a consensus was quickly reached about it being too cumbersome and too much of an inconvenience:

> *"I hate that! It's incredibly lame, I don't understand how it works!"*

followed up from a different participant with:

> *"The user has to consciously walk up to and seek out the information, that can avoid spam. But yes, QR is a bit of a hassle. It's a bit cumbersome to use. Because you need a dedicated application. It's way easier to simply put a card up to something than to bring up an application and hold it completely still and hover it above a code."*

Two of the participants had experienced that their cracked screen and insufficient screen light on different occasions had hindered QR readers to read the QR codes on their phones. In the case of insufficient screen light they had to increase the light but, even though a relatively quick task, it was seen as inconvenient. One of the participants mentioned that his cracked phone screen almost prevented him from entering a concert as the reader took too long to properly read the code.

6.3.2 Automation of tasks

During the introduction of the focus group I mentioned that my research was focused on people's trust in technology and automation. Immediately upon hearing this, one of the participants said

> *"I don't trust technology, I believe it's taking over the world. One can already hack into a person's pacemaker and kill them"*

[3] www.google.com/wallet
[4] www.apple.com/apple-pay

fearing that people were becoming too dependent on technology and that we might be rendered helpless if the technology were to someday fail. The participant with the biggest faith in technology of all the members of the focus group stated that he believed Bluetooth beacons were the next step in payment and authentication:

> *"Then you can simply walk around and the phone takes care of everything".*

When again describing the Android application I developed and tested with users, the initial question of one participant was about what would happen if you went on the metro, but then immediately changed your mind. She stated that she would not approve of such a system, as it would purchase a ticket even though she ended up not needing it:

> *"What if I you go the station and discover that you're not going to take the metro after all. Then you have to go back, and haven't taken the metro. That's a bit stupid. I would have hated that system."*

Regardless of the functionality of cancellation, all four participants were afraid that such a system would make purchases for them that they did not want to be made. As a response to this concern I mentioned that one earlier participant suggested that the phone asks you to confirm a payment instead of automatically making it for you. The four participants immediately agreed that push notifications would be nice, one of them stating:

> *"I fully agree. That [the phone] doesn't say 'Now we withdraw some money from your card just because you are in this area'. If you get a notification where you have to confirm, it's still incredibly easy. You don't have to open up an app and do lots of stuff".*

The participant with the biggest faith in and enthusiasm for technology said that it would be nice to be able to "live seamlessly" and let a computer do all the transactions for you. However, he was not convinced that the technology was able to make life entirely seamless just yet. Hence, he liked the functionality of the confirmation button:

> *"I like that. Quite simply push notifications. Push to the front of the screen and ask 'Do you want to pay?'"*

After some discussion on the topic of beacons, one participant realized that beacons placed everywhere around the city, not necessarily just for the use on public transport, could be a huge source of radiation. This lead the two participants who were afraid of radiation from the Bluetooth on their phones to feel not all that happy about many beacons placed around their town. They were afraid that even though they disabled Bluetooth on their phones they would still receive radiation from the beacons in their immediate surroundings. A concern of all the focus group participants was that of spam. They were afraid they would be bombarded, not necessarily with automatic payments, but also with advertisements, offers and information. A limit on how close you have to be before you will be notified of these advertisements was seen by the participants as fairly important in lowering spam:

> *"Let's say that you get a push notification when you're close to the metro station, then you would need some sort of limit though. Let's say you're in Oslo and you walk down a street. You can't all of a sudden get 70 push notifications for museums and stuff".*

As an addition to this, one participant mentioned that such services should be opt-in, not opt-out, meaning you would have to actively register your name or phone to receive such notifications. He felt this would help limit the amount of spam you would receive simply from walking down a street or being in a particular area. He pointed to this area as a place where parts of QR codes were good: that you have to actively seek out and interact with the codes yourself. Still, however, he felt that the many cons of QR codes outweighed this pro.

> *"If you're taking the bus you have to be registered at [Ruter] in order to receive such push notifications. Because after a while the town will be littered with beacons."*

After listing both negatives and some positives with the use of beacons to automatize tasks, the participants were asked whether or not the automatic task involving money would make it more deterring. They all quickly and decidedly answered yes, two of them seeming stupefied that a question with such an "obvious" answer was even asked. This deterring factor came from a potentially bigger loss if the system were to be hacked, as well as a fear of losing control of their money. One participant stating

> *"I would never use anything that pays for me. I want to control my own money. I want to always see how much I have and what goes out."*

On this topic the same participant who liked the thought of eventually being able to live "seamlessly" through automation, said that people who struggle with their personal finances could end up struggling even more because of a lowered threshold to spend money:

> *"It becomes such a natural part of your everyday life that you don't even think about it. Money is withdrawn from your account here and there. You walk into the coffee shop and it asks you if you want the same coffee as yesterday, you press 'OK' and you don't even think about it. And all of a sudden you're broke".*

Towards the end of the focus group, the participant who had been adamant that automation could lead us to being too dependent on technology said that she was afraid automating too many tasks could come at the expense of losing natural human intercommunication. Communication with the bus driver or the cashier at the coffee shop was for her a valued experience that she feared would disappear if the world became too automated.

> *"I think that it's very good when we get techniques and technology that makes it easier to shop and pay and so on, but not when it comes at the expense of how we interact with each other. I'm still supposed to be able to walk into a store and an employee should greet me and ask 'How can I help you?'"*

She was then asked whether or not she used the self check-in machines at the airport or self-serve kiosks at restaurants, or if she went to a person behind a counter, to which she answered that she never used the self check-in machines if she was traveling alone. When traveling with someone else, however, she was never in charge, hence the others usually chose to use the self check-in machines. The interaction with the person behind the counter made her more confident in that nothing was wrong and that she was checked in successfully.

> *"The other week we checked in a bag ourselves, but we forgot to check how much it weighed, and it doesn't tell you that. But if you go to the counter they tell you if it weighs more than 20 kg. Then I just sat there thinking 'Crap, are they gonna send me a huge bill?'"*

She felt the feedback from the self check-in machines were not good enough, and that she lost some control when not checking in through a person behind a counter. A different participant immediately said:

> *"Oh, that's so awesome, I love that! Instead of talking to a person you just type it in yourself. It's so awesome. When I'm at the airport I can just walk up to the machine and type in my info, walk to the bag drop-off and do it myself. The only place I have to communicate with people is at the gate."*

The two other participants also used self check-in machines every time they went to the airport as they felt it machine gave sufficient feedback if something was wrong.

Chapter 7

Discussion

This chapter discusses the research questions based on the theory and my findings. The research questions of this thesis are related to the use and understanding of ubiquitous systems, specifically context-aware automatic payment systems. A system might not be understood or used to its full potential or as it was intended by the creators, because the users have a different mental model than what the creators envisioned and what the conceptual model the system was built upon. A poor understanding of or negative reaction to such context-aware systems might show that gradually easing the users into these new concepts might prove effective, allowing their mental models to develop over time. To reiterate, the research questions of this thesis are:

1. How do the users' mental models of Bluetooth affect the use of beacon technology for automated context-aware payment?

2. What, if any, are the issues users face with regards to perceived control when using automated context-aware payment?

3. In what way does testing a ubiquitous computing system in the field result in different findings than testing in a lab?

The research questions will in this chapter be discussed one by one. It is important to note that many of the findings of this thesis are the users' statements on how they would have acted and explaining their expectations of the system. A large part of the discussion is about what might affect the users' trust and sense of control. If they only trust the system to a limited extent they won't even bother trying it out, making their thoughts and expectations of the system an important topic to research and discuss. If they don't want to pick it up in the first place, it's difficult to change their views. Spiekermann (2008) mentions this when saying that a challenge in ubiquitous computing is that a prototype needs to be fairly advanced in order to test it with users. Hence, she argues that prospective findings are of great importance in ubiquitous computing, not just retrospective. I would say that the prototype I developed and tested was close to as simple as a ubiquitous computing prototype can become, and that it helped the users understand the premise and concept of the system, giving them a reference point during the interviews to discuss their thoughts, concerns and expectations.

7.1 The user's mental model of Bluetooth

During the interviews and focus group participants talked about how they had previously used Bluetooth, what their thoughts about and experiences with it were, and their reasoning for either using it or not using it. This gave me a good understanding of their mental model of Bluetooth, which in many cases seemed to consist of some aspects of uncertainty.

Even though classic Bluetooth and Bluetooth low energy (BLE) are not the same technology, most users have little reason to believe otherwise. Additionally, a phone doesn't normally say if it only has BLE scanning enabled or classic Bluetooth as well.

7.1.1 Power consumption

The participants were generally afraid that Bluetooth consumed a substantial amount of power, or at least "felt" like it did. They were fairly uncertain, but seemed to remember Bluetooth's power consumption being a serious issue several years ago. This might have changed since then, they were unsure. For example, one participant stated:

> *"I also feel like it consumes power, I don't know how much it consumes, but at least it used to, I don't know if it does any longer. I don't use it for anything, so there's no point having it enabled."*

As Gentner and Stevens (1983) say, if the mental model consists of uncertainties, the user will normally approach the system or scenario with caution. Then Gentner and Stevens say, through interaction and experience with the system, their mental models evolve and their uncertainties might disappear. As the users are uncertain about Bluetooth's power consumption, a desire to try it out and see for themselves might emerge, allowing them to rid their mental models of this uncertainty. However, if they are skeptical when trying this new technology, a confirmation bias that Nickerson (1998) talks about can take place. For instance, if for some other reason the phone starts consuming more power around the time the user tries the application, he might believe and convince himself that it is because of only this very application. A feeling of "I was right!" can be induced. Or perhaps the user had not paid much attention to the battery consumption on his smartphone before, and even though the consumption barely increased, he might think that the battery not lasting as long as he'd like is the cause of this newly downloaded Bluetooth based application.

As can be seen in the study conducted by Aislelabs (2014), BLE scans and operations generally consume a minuscule amount of power. This is especially true if the developers implement a scanning strategy similar to what GSMA (2014, p. 41) suggests. The newer the phone model is, the less power is generally consumed, showing an ongoing improvement in BLE chip technologies. The amount of nearby beacons, however, can have an impact on the battery consumption (Aislelabs, 2014). This can add to the issue of spam and radiation exposure if beacons are to be scattered densely throughout the environment. The participant who himself was writing a thesis on Bluetooth beacons had read several articles showing that Bluetooth consume little power, but he still had trouble accepting its low power consumption. As a result he always disabled Bluetooth. This can show that the mental models are not so easily changed. This might prove to be a substantial mental model obstacle when releasing such applications and trying to get people to use BLE for location tracking. Users not being convinced that enabling Bluetooth on their phones won't consume all their power is important when introducing BLE based tracking. After all, having Bluetooth enabled is a prerequisite for this kind of tracking. If the users don't feel confident in having Bluetooth enabled, the application may not be used at all.

The participants' views and feelings on Bluetooth were generally not of a negative nature, despite their uncertainty regarding the power consumption. Many of them concluded that the reason they did not have Bluetooth enabled at all times was mostly because they did not really use it for anything:

> *"It's really just that I don't have any need for Bluetooth. Why would I have it enabled if I don't use it?"*

And again, when the participants were uncertain about Bluetooth's battery consumption, they seemed to be well aware of it and of the fact that the battery consumption might not be as substantial as they felt it used to be.

This power consumption uncertainty I would argue can be changed through what Gentner and Stevens (1983) call indirect exposure. Word of mouth or reading about the system on a poster on the metro might slowly evolve the users' mental models into understanding and believing that BLE indeed has a fairly low power consumption. Again, especially when the application implements techniques such as suggested by GSMA (2014, p. 41). Information about and endorsement of the system from Ruter was after all mentioned as a way in which the users could establish trust and confidence in it, which I discuss in section 7.2.

7.1.2 Location estimation

Across all the user tests as well as the focus group, location tracking was part of only one of the eleven participants mental model of Bluetooth. However, at the time this one participant was writing a master's thesis about the use of Bluetooth beacons in a museum setting, giving him expert-level experience and knowledge in the area. He, however, admitted that he had not really heard about beacons before he started on his thesis. Two other participants knew that beacons could be used for advertising. They did not, however, know how it worked or that one could measure one's distance from a beacon. As such, location tracking was not something the participants thought Bluetooth was capable of. When they saw the application working and that a ticket was indeed purchased, however, they became pleasantly surprised, one of them exclaiming "It actually works!". Location tracking and estimation had not been a part of what any participant had used Bluetooth for in the past, they had only previously used it to pair devices to send files or listen to music. This shows that their mental model of Bluetooth did not include the peripheral, broadcasting or observer modes of BLE devices (Lindh, 2015), only the central mode. They associated Bluetooth communication with paired two-way communication, not an broadcasting one-way communication platform.

This may affect whether or not people feel a desire to try such a system, as they might not believe it's possible to accurately estimate their location just through Bluetooth.

This should be taken into account when developing systems in general based on BLE location tracking, as too much new and innovative functionality at once can overwhelm the users. As Kuniavsky (2010) says, focusing on the core functionality is important in innovative applications, and preferably building the rest of the system on experiences that are similar to what the users have used before. Too drastic innovation all at once can quickly cause confusion as the users' mental models need to evolve. I will elaborate on this in the next section, along with the participants' perceived control and trust as I discuss the second research question.

7.2 Automation and perceived control

This section discusses the second research question, relating to how the users perceive their control over the application. It quickly became apparent that the participants felt uneasy about the fully automatic functionality, and that this uneasiness was enhanced by the fact that money was involved.

7.2.1 Trusting the technology

In the user tests, several participants seemed to struggle to understand that the application was going to automatically purchase a ticket for them, and that they did not have to directly interact with the application. Several participants had to ask more than twice if they had understood the task correctly when they were thinking they should do nothing but keep the phone in their hands or pocket while walking to the station. This was a new concept for them, not fitting their mental model of Bluetooth or of ticket systems in general. One participant asked if he had to order the ticket beforehand, showing that he figured that interaction was needed from him at least at some level and some point in time.

Some participants were afraid the application would purchase too few tickets, other afraid it would purchase too many. If they knew that the application was released by Ruter, however, and not a third party developer, they would feel more confident using it. As one participant said:

> *"If it's supported by Ruter it should work. Because I expect Ruter to have some sort of quality."*

As they were uncertain as to whether the Bluetooth location estimation technology would work properly, the legitimacy of a company such as Ruter would provide some confidence. Firstly, they felt that Ruter would ensure the quality of the application. They knew that the RuterBillett application already worked, and implementing the functionality into that already existing application were by a few said to further increase their confidence in it. Secondly, if the application did something wrong at some point, Ruter could provide a better customer support service than a third party. These aspects wouldn't necessarily give the user more control, but a better *perception* of control, as well as a sense of security. As the participants said, if Ruter supported the application, Ruter would be reliable if something went wrong. One participant expressed this by stating:

> *" [...]If something is wrong, I don't know what it could have been, either that it doesn't purchase a ticket or charges me twice, then it can be fixed by Ruter. If the application was not developed by Ruter, the company could say 'Thats not our responsibility, contact [Ruter]'".*

Four participants said they would have to manually check that their ticket indeed had been purchased, at least for the first few weeks, as their trust in the technology not would have been big enough even if Ruter themselves had released the application. The uncertainty residing in their mental models of Bluetooth could cause them to lose all trust entirely from a single slip-up, as mentioned by one participant. Uncertainty and a "seems to work" attitude, like Gentner and Stevens (1983) talks about, seemed insufficient as the application was dealing with money. Checking that the ticket was purchased would give them a sense of control, but at the same time part of the autonomy would be gone, which will be further discussed in the next section.

7.2.2 Being in control

After the user tests, the participants all understood what requirement needed to be fulfilled before a ticket was purchased - being on the metro station. Although, as mentioned above, some were more afraid than others that the location estimation wouldn't be completely accurate. When asking if money being involved made the loss of control more intimidating, all participants responded with a resounding "Yes!" - without hesitating. It seemed some participants thought this was an unnecessary question with an "obvious" answer.

When I mentioned the option of passive context-awareness, the participants responded positively. This suggestion came from one of the participants saying:

> *"No, I wouldn't have felt safe using it. I wouldn't have managed to cancel the tickets in time. Maybe I didn't feel the phone*

> *vibrate. [...]I'd rather have to press 'OK' instead of having to press 'Cancel'. 'OK' if I want a ticket."*

After this they talked about automation as if the passive context-awareness and the application asking for confirmation was given, indicating that this was their preferred level of control. However, two participants in particular figured that this would defeat the purpose of the application and its complete automation. As one participant said in response to the passive context-awareness:

> *"Then you have to do more yourself [...]. A part of the point is gone. The point [of the application] is that you can forget about it and still have a ticket."*

Some of the purpose of the automation would have disappeared. The user would not have been able to completely forget to purchase a ticket, and a passive context-aware system will interrupt the user's current task to ask for confirmation, creating a context switch. However, a large part of the process is still left to the application as it's still automated on some level. Gudymenko and Borcea-Pfitzmann (2011) mention a "transition of computing power to the background". I found that the word "transition" here is important. It explains well how the participants generally felt that automated systems, especially those dealing with money, needed to be handled. They could not put their trust entirely in a fully automated application, but needed first to be able to make the final call, as Bellotti and Edwards (2001) call it. Evolving the mode and level of interaction slowly towards full automation will be necessary, as it will help users still feel in control. A quick revolution of automation can become too intimidating, especially when involving money. The preference-based context-awareness of Hardian et al. (2006) allows the users to personalize the behavior of the application, deciding themselves how it should behave in certain situations and how much control they would like to have. This was also suggested by one of the participants:

> *"At least ask. Or that you can choose yourself, but as a standard setting it asks you [before it purchases]. Custom settings. Are we not in 2016?"*

Starting out with a passively context-aware system, they can move on to an actively context-aware system if and when they see fit. As two participants mentioned, being able to customize the level of control would allow each user to give the application the amount of control they felt comfortable with, one of them feeling that preference-based and customizable behavior was only natural in the technological world of 2016. Having the system behave in a passive context-aware way by default, the users can then decide to make it actively context-aware if they please. Barkhuus and Dey (2003) said that users prefer to use more automated systems as long as the usefulness is greater than the cost of losing control. This cost may seem bigger when not used to and confident in the technology, and the cost is literal when the application deals with monetary operations. I would argue that the need for a preference-based control scheme is supported by van der Heijden (2003), when he says that the desired amount of control is different from one person to another.

This passive context-awareness was preferred to a cancellation button, as some were afraid they wouldn't be able to cancel the ticket before the one minute timer had run out. This is in line with the thinking of Bellotti and Edwards (2001), where a slight doubt in the user's outcome would need an effective way to correct the actions taken, while a significant doubt should prompt the user for a confirmation rather than take an automatic action. Involving money in the process increases the consequences of making the wrong decision, and should increase the doubt accordingly. This is supported by Kindberg and Fox (2002) when they talk about the act of establishing a clear "semantic Rubicon" - a clear definition of what is the user's responsibility and what is the system's responsibility. When ambiguity arises, Kjeldskov and Fox say that the user should resolve the situation, which would be the case of a passive context-aware system. Even if the system would interpret a situation as unambiguous, providing the user with the ultimate decision would help induce a sense of control. After all, the user might not trust the system, and underestimate its abilities to correctly establish the user's context.

Kuniavsky (2010) says that the more complex an innovation is, the more difficult it will be for the users to grasp and understand it. They will have nothing to compare and relate it to. Experiences with similar systems are important when users establish mental models of a new system. I see this as yet another argument for starting out with a passive context-aware system, where the users can get accustomed to the experience. The act of automating systems through context-awareness is not yet pervasive, especially not in monetary operations. If first experiencing a passive context-aware system, the users will perceive control while some level of automation is introduced. Then, later, one can step up the level of automation and complexity of the innovation, gradually evolving the users' mental models on, trust in and perceived control over automated systems.

One problem that can occur from a prevalent use of passive context-awareness, however, is that of sloppiness from the software developer side in estimating the user's location. If no users have a fully automatic active context-aware application, there might not be a great need from the developer's point of view to invest time into improving and perfecting the location estimation. After all, the users make the final call each time. This can lead to a self-fulfilling prophecy: the developers feel no need to perfect the location estimation when no one is using active context-awareness, hence users don't feel confident moving to active context-awareness as the system is not optimal, and as a result the developers confirm that perfecting the location estimation is unnecessary. In the long term this might hurt the area of ubiquitous and context-aware computing, as the transition to full active context-awareness is slowed down or inhibited. The developers need to be aware of the fact that users will not move to and consistently use an active context-awareness before the system is working properly and with a minuscule amount of errors.

The participants in the focus group were especially worried about spam when talking about beacons in general. They argued that automation systems based on beacons should be based on an opt-in model where applications would not execute actions based on the user's context unless he actively gave it permission, such as by downloading the application. They needed to have ultimate control over the application, not having to take part in it at all if they did not want to, and easily being able to opt out of it after having opted in. As Eddystone scanning is being implemented into mobile Chrome browsers, the threshold for receiving these beacon advertisements are lowered, as downloading a specific application is not required. Not opting in to such functionality is, for now, an option, and disabling Bluetooth even though you have opted in will prevent the scanning from happening. However, a concern for phones still enabling Bluetooth without letting the user know was mentioned by The Norwegian Data Protection Authority, and according to them has happened before (Årnes, 2016). As of now, the ultimate control to opt in and out of these Bluetooth beacon tracking systems lies with the user. However, this might not be the case in a not too distant future.

As can be seen from the discussion of this research question, the participants generally didn't feel confident in giving a smartphone application complete control of their money. Even if it were released by Ruter, they would only feel truly confident using it if there was some passive context-awareness aspect to it, where they were asked if they would want to purchase a ticket instead of the application automatically purchasing it for them. This holds true at least when adjusting to and getting to know the technology, making sure that it really works. This would help the transition into a world of active context-awareness. It should be noted, though, that when not dealing with money, an active context-aware application would probably have been received better than it was in this case, as all participants clearly stated that money being involved made the automation more intimidating.

The next section discusses the third and last research question, regarding the differences of testing a ubiquitous system in and out of its intended use context.

7.3 Testing in or out of context

This section discusses the third research question, relating to the benefits of testing a ubiquitous system in a lab or in-situ with a realistic context. It needs to be stressed that only three participants tested the system in the lab, while four tested it in-situ on the metro, resulting in a fairly low sample size.

I made three main observations as differences between the lab and in-situ tests: (1) the in-situ participants were the most pleasantly surprised by the autonomy, (2) the semi-structured interviews with them flowed more smoothly, and (3) only the in-situ participants were concerned about the legitimacy of the tickets. This last observation is similar to one seen by Kjeldskov et al. (2004). They saw that the participants in field tests expressed a concern about data validity, which none of their lab test participants expressed. This shows a better level of immersion and that a more realistic setting can bring forth a different set of issues from the participants. Like Kjeldskov and Skov (2014) said, lab tests tend to find more issues related to interface design, while field tests can be very effective for developing and testing the more general concept of a system, which I found to indeed be true.

The participants of the focus group had a more negative outlook on the entire automation process than the participants of the user tests, both in the lab and in-situ. They were more afraid of spam, radiation and the loss of control than those in the user tests. This may stem from the fact that they lacked hands-on experience with the system and didn't see that it indeed does work. This is in line with what Rogers et al. (2007) mean when arguing that in the field of ubiquitous computing, with such new and innovative interactions, being able to test the system, especially in-situ, is important. One of my field testers even stated that he did not really think this could work until he tested it and saw it working in action. On a related note, the in-situ participants were more amazed and pleasantly surprised when the ticket was purchased, than the participants of the lab tests This yet shows an increased level of immersion and realism when testing in a realistic context. The in-situ participants got to see that the system can actually work in the real world, not just in the controlled conditions of the lab. Spiekermann (2008) mentions that the users that perceive the most control over a system also have the most positive attitudes towards it. If we turn this around, we can argue that the users with the most positive attitudes towards the system perceived the most control over it. Based on this and the fact that the in-situ testers had the most positive attitudes towards the system, I argue that the in-situ tests and hands-on experience gave the most realistic representation of the system and the best perception of control. This argument is strengthened when we see that the focus group participants were the most negative of all participants, having had no hands-on experience at all. I found that the more hands-on and realistic the participants' experiences were with the system, the more positive their attitudes toward it was.

The in-situ testers on the metro asked more questions than the testers in the lab, and were generally more talkative during the semi-structured interviews. This observation is similar to one Kaikkonen et al. (2005) made about users being more capable of expressing themselves and their experiences, opinions and concerns about the concept if they had tested it in a field test rather than a lab test. In the case of my thesis, an application that demands very little attention of the user was tested. As both the technology and interaction type is rather new, a first-hand realistic experience with the system resulted in interviews with less hypothetical thinking. They have seen for themselves that it actually might work.

The act of recording the in-situ tests on video has been mentioned as a difficult task. Recording setups such as a camera attached to the mobile phone the participant is using or on the participant himself with equipment carried in a backpack (Kaikkonen et al., 2005) can disrupt the participant's work and ruin immersion. Additionally, a camera operator and an observer following a participant in the field can cause for unrealistic behavior from the people the participant meet. In the case of Kjeldskov and Stage (2004), people moved away from the testing "team" as they walked down a pedestrian street, resulting in a less realistic test situation. I did not record any video during my tests, but will still like to raise a point regarding video capture in field tests. With the improvement of technology and invention of even smaller cameras, these issues are starting to fade away. With the advent of small cameras attached to a person's chest or shoulder, such as the GoPro[1], video recordings of field studies can be conducted without the use of a dedicated person to operate the camera, and without attachments to the phone

[1] www.gopro.com

that might obstruct the participant's use. Being able to record the participants' surroundings is in some cases necessary. With the development of 360 degree cameras the task of capturing the surroundings can be completed in its entirety by a single camera. And with virtual reality technology rapidly being available to consumers in 2016, the 360 degree videos the 360 degree cameras produce can help the researchers being "taken back to the scene of the research" - at least virtually. This technology is still in its infancy, however, recording the participant and his surroundings during field tests is less of a problem today than it was only a few years ago.

Chapter 8

Conclusion

Through this thesis I have looked at the control users perceive when automating tasks through Bluetooth beacons, and how comfortable they are giving control of monetary operations to a smartphone application. Below I summarize the research questions and their accompanying discussions one at a time. Lastly, I talk about how others can continue and further my work.

8.1 Mental models of Bluetooth affecting expectations

The first research question regards how the users see and experience Bluetooth, which might affect their willingness to or behavior during use of systems based on Bluetooth beacons.

Research question: How do the users' mental models of Bluetooth affect the use of beacon technology for automated context-aware payment?

Very few participants used Bluetooth on a regular basis, and only one of them had it enabled at all times. All participants expect the one who always had Bluetooth enabled said that they feared Bluetooth consumed a substantial amount of power. Most of them concluded that the main reason for them not having Bluetooth enabled was the fact that they just did not use it for anything. They saw no point in having Bluetooth enabled when they did not use. Regarding the power consumption, the general stance of the participants was that Bluetooth might or might not consume much power - they were not sure. They remembered Bluetooth consuming a substantial amount of power some years ago, but admitted the technology might have improved since then. The fact that Bluetooth could be used for location estimation was a new concept to most participants, however, they had no trouble believing it to be true when I mentioned it. Because of their lack of knowledge and their uncertainties regarding Bluetooth, their mental models seemed to be open for change and learning. As a result, I do not deem the users' mental models of Bluetooth to affect their use of or chances of approaching Bluetooth beacon based systems in a substantial way. The gaps of their mental models caused by uncertainties can quickly be filled from experience and exposure to the system, or even through in-direct exposure such as marketing or word of mouth. The more skeptical to the technology they are, however, the less willing they can be to change their mental models. When deploying Bluetooth beacon based systems one need to keep in mind that users generally don't know all that much about the Bluetooth technology, and that power consumption is a concern for many.

8.2 Passive context-awareness as the first step

The second research question focuses the users' perceived control and trust in using the tested prototype - a smartphone application that purchases a metro ticket when the user steps on the station.

Research question: What, if any, are the issues users face with regards to perceived control when using automated context-aware payment?

Automation was by the participants seen as more intimidating and the loss of control more serious when involving money. The loss of control would be too big for the participants to put their full trust in an actively context-aware application handling their money. At least without any form of transition and time to make sure the system indeed works consistently and as it should. When moving towards the world of ubiquitous computing, one should be mindful when deciding on the level of context-awareness, as too great a loss of control will turn the users away from your system. Especially for monetary operations, passive context-awareness where the user is prompted to accept an action should be enabled by default. The users can then, through the application's settings function, decide to make the system actively context-aware and let it behave as it pleases.

8.3 Grasping new concepts through field tests

The third research question regards the effects of testing a ubiquitous system in an in-situ field test as opposed to testing in the lab.

Research question: In what way does testing a ubiquitous computing system in the field result in different findings than testing in a lab?

The focus group was not strictly a type of user test, however, it portrays a research scenario in which the participants had no hands-on experience with the prototype whatsoever. The least amount of experience with and hence immersion of the prototype was found in the focus group. Consequently, the focus group participants expressed the most concerns regarding the system and the most negative attitude towards it. During the non-realistic user tests in the lab, the participants got to try the prototype, but not out in the real world. They were more positive to the concept than the focus group participants, and still raised concerns. The participants of the field studies contributed the most to the post-test interviews and these interviews flowed better than the interviews after the lab tests. Additionally, the participants of the field tests had the most positive view towards the concept of the prototype, and were the most pleasantly surprised during testing.

Based on my findings here I conclude that interaction with a prototype is essential when designing ubiquitous systems, even though the prototype is not fully functional. Testing the prototype in the field gives the participants a more realistic ground when discussing the system, and allowing them to see the system as more than an isolated system in a lab. It should be noted, however, that prospective findings and data on the users' intentions to approach the system still can be gathered with little to no hands-on experience with a prototype.

8.4 Future work

Iterating on and implementing the design solutions proposed in this thesis followed by a test in association with a public transportation company, would be a good direction to take this research in next. Being able to deploy beacons at the stations the users usually use and working out a deal with Ruter where the tickets purchased with the prototype are legitimate, would allow the users to use the system on a bigger scale. Additionally, one can then avoid speculation about aspects of use such as the time it would take for users to trust the application, or whether or not a passive context-aware solution would be too obtrusive. Abowd and Mynatt (2000) claim that ubiquitous systems need to be deployed in a realistic context to be tested properly. In my research, the scaling dimensions Abowd and Mynatt mention - device, space, people, and time - are not all properly tested. These scalings could be put to a more realistic test with a deployed pilot system in cooperation with Ruter. The data gathered from such a study could be of a mixed nature - quantitative data complemented by qualitative interviews and focus groups. Kjeldskov and Skov (2014) mention the use of more longitudinal studies of mobile systems in order to better grasp the users' experiences. As discussed in this thesis, the application to be used in further testing should preferably implement a passive context-awareness where the users are reminded and prompted to purchase a ticket with a single click when entering a station. However, they should have the option to make the system fully automatic and actively context-aware. For research purposes, the application can track how long it takes for the users to switch to fully automatic purchase. This can help facilitate qualitative discussion or simply be collected as quantitative data. Furthering the research like this will help gather retrospective data where this thesis only found prospective data, and will further improve the realism of the tests.

Bibliography

Abowd, G. D. & Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction*, *7*(1), 29–58.

Aislelabs. (2014). Ibeacon battery drain on apple vs android: a technical report. Retrieved February 23, 2016, from http://www.aislelabs.com/reports/ibeacon-battery-drain-iphones/

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179–211.

Apple. (2014). *Getting started with ibeacon*. Retrieved from https://developer.apple.com/ibeacon/Getting-Started-with-iBeacon.pdf

Araya, A. A. (1995). Questioning ubiquitous computing. In *Proceedings of the 1995 acm 23rd annual conference on computer science* (pp. 230–237). ACM.

Argyris, C., Putnam, R., & Smith, D. M. (1985). *Action science*. Jossey-Bass social and behavioral science series. Jossey-Bass.

Årnes, A. (2016, January). *Afsecurity seminar: bluetooth beacon privacy*. Presentation held at the University of Oslo by the Director of Tehcnology at The Norwegian Data Protection Authority.

Bahl, P. & Padmanabhan, V. N. (2000). Radar: an in-building rf-based user location and tracking system. In *Infocom 2000. nineteenth annual joint conference of the ieee computer and communications societies. proceedings. ieee* (Vol. 2, pp. 775–784). Ieee.

Barkhuus, L. & Dey, A. (2003). Is context-aware computing taking control away from the user? three levels of interactivity examined. In *Ubicomp 2003: ubiquitous computing* (pp. 149–156). Springer.

Baxter, P. & Jack, S. (2008). Qualitative case study methodology: study design and implementation for novice researchers. *The qualitative report, 13*(4), 544–559.

Bellotti, V. & Edwards, K. (2001). Intelligibility and accountability: human considerations in context-aware systems. *Human–Computer Interaction, 16*(2-4), 193–212. doi:10.1207/S15327051HCI16234_05. eprint: http://www.tandfonline.com/doi/pdf/10.1207/S15327051HCI16234_05

Buie, E. (1999, March). Hci standards: a mixed blessing. *interactions, 6*(2), 36–42. doi:10.1145/296165.296177

Buxton, B. (2013, March). Bill buxton at techfest 2013: designing for ubiquitous computing. [video file - microsoft research]. Retrieved November 11, 2015, from https://youtu.be/ZQJIwjlaPCQ

Chen, G. & Kotz, D. (2000). *A survey of context-aware mobile computing research*. Dartmouth College.

Cook, B., Buckberry, G., Scowcroft, I., Mitchell, J., & Allen, T. (2005). Indoor location using trilateration characteristics. In *Proc. london communications symposium* (pp. 147–150). Citeseer.

Craik, K. (1967). *The nature of explanation*. Philosophy: science CAM. Cambridge University Press.

Crang, M. & Cook, I. (2007). *Doing ethnographies*. SAGE Publications Ltd. doi:http://dx.doi.org/10.4135/9781849208949

Datatilsynet. (2014). Årsmelding for 2014.

Datatilsynet. (2015). Det store datakappløpet.

Dourish, P. (2004). *Where the action is: the foundations of embodied interaction*. MIT press.

Faragher, R. & Harle, R. (2014). An analysis of the accuracy of bluetooth low energy for indoor positioning applications. *Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+'14).*

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry, 12*(2), 219–245.

Friedewald, M. & Raabe, O. (2011). Ubiquitous computing: an overview of technology impacts. *Telematics and Informatics*, *28*(2), 55–65.

Furht, B. (2011). *Handbook of augmented reality*. Springer Science & Business Media.

Gentner, D. & Stevens, A. (1983). *Mental models*. Cognitive Science - Lawrence Erlbaum Associates. Taylor & Francis. Retrieved from https://books.google.no/books?id=QFI0SvbieOcC

Google. (2015). Eddystone protocol specification. Retrieved from https://github.com/google/eddystone/blob/master/protocol-specification.md

Group, N. N. (2014). Turn user goals into task scenarios for usability testing. Retrieved November 5, 2015, from https://www.nngroup.com/articles/task-scenarios-usability-testing/

GSMA. (2014). A guide to bluetooth beacons. Retrieved from http://www.gsma.com/digitalcommerce/wp-content/uploads/2013/10/A-guide-to-BLE-beacons-FINAL-18-Sept-14.pdf

Gudymenko, I. & Borcea-Pfitzmann, K. (2011). Privacy in ubiquitous computing. In *Interconnecting smart objects with the internet workshop, prague* (p. 60). Citeseer.

Hardian, B., Indulska, J., & Henricksen, K. (2006). Balancing autonomy and user control in context-aware systems-a survey. In *Pervasive computing and communications workshops, 2006. percom workshops 2006. fourth annual ieee international conference on* (6–pp). IEEE.

Harrison, S., Tatar, D., & Sengers, P. (2007). The three paradigms of hci. *alt. chi*.

Hassenzahl, M. & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, *25*(2), 91–97.

Houde, S. & Hill, C. (1997). What do prototypes prototype. *Handbook of human-computer interaction*, *2*, 367–381.

HP. (2002). *Wi-fi and bluetooth - interference issues*.

James Rumbaugh, G. B., Ivar Jacobson. (2004). *Unified modeling language reference manual, the (2nd edition)*. Pearson.

Jr, W. S. M. (2007). Determination of a position using approximate distances and trilateration. *Colorado School of Mines*.

Julsrud, T. E., Denstadli, J. M., Herstad, J., Hjalmarsson, A., & Li, Y. (2014). *Mobile applikasjoner underveis: hva ønsker de reisende?*

Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2005). Usability testing of mobile applications: a comparison between laboratory and field testing. *Journal of Usability studies*, *1*(1), 4–16.

Khandkar, S. H. (2009). Open coding. *University of Calgary.*

Kindberg, T. & Fox, A. (2002, January). System software for ubiquitous computing. *IEEE Pervasive Computing*, *1*(1), 70–81. doi:10.1109/MPRV.2002.993146

Kitzinger, J. (1994). The methodology of focus groups: the importance of interaction between research participants. *Sociology of health & illness*, *16*(1), 103–121.

Kjeldskov, J. & Skov, M. B. (2014). Was it worth the hassle?: ten years of mobile hci research discussions on lab and field evaluations. In *Proceedings of the 16th international conference on human-computer interaction with mobile devices & services* (pp. 43–52). MobileHCI '14. Toronto, ON, Canada: ACM. doi:10.1145/2628363.2628398

Kjeldskov, J., Skov, M. B., Als, B. S., & Höegh, R. T. (2004). Is it worth the hassle? exploring the added value of evaluating the usability of context-aware mobile systems in the field. In *Mobile human-computer interaction-mobilehci 2004* (pp. 61–73). Springer.

Kjeldskov, J. & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *International journal of human-computer studies*, *60*(5), 599–620.

Krishnan, P., Krishnakumar, A., Ju, W.-H., Mallows, C., & Gamt, S. (2004). A system for lease: location estimation assisted by stationary emitters for indoor rf wireless networks. In *Infocom 2004. twenty-third annualjoint conference of the ieee computer and communications societies* (Vol. 2, pp. 1001–1011). IEEE.

Kuniavsky, M. (2010). *Smart things: ubiquitous computing user experience design.*

Langheinrich, M., Coroama, V., Bohn, J., & Rohs, M. (2002). As we may live: real-world implications of ubiquitous computing. *ETH Zurich.*

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 719–728). ACM.

Lindh, J. (2015). *Bluetooth low energy beacons*. Texas Instruments.

Liu, L. & Khooshabeh, P. (2003). Paper or interactive?: a study of prototyping techniques for ubiquitous computing environments. In *Chi '03 extended abstracts on human factors in computing systems* (pp. 1030–1031). CHI EA '03. Ft. Lauderdale, Florida, USA: ACM. doi:10.1145/765891.766132

Lucky, R. (1999). Everything will be connected to everything else. connections. *IEEE Spectrum, March*, 19.

Mantei, M. M. & Teorey, T. J. (1988, April). Cost/benefit analysis for incorporating human factors in the software lifecycle. *Commun. ACM*, *31*(4), 428–439. doi:10.1145/42404.42408

Myers, B. A. (1998, March). A brief history of human-computer interaction technology. *interactions*, *5*(2), 44–54. doi:10.1145/274430.274436

Myers, M. D. et al. (1997). Qualitative research in information systems. *Management Information Systems Quarterly*, *21*(2), 241–242. Retrieved from http://www.qual.auckland.ac.nz/

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Nielsen, C. M., Overgaard, M., Pedersen, M. B., Stage, J., & Stenild, S. (2006). It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. In *Proceedings of the 4th nordic conference on human-computer interaction: changing roles* (pp. 272–280). NordiCHI '06. Oslo, Norway: ACM. doi:10.1145/1182475.1182504

Bluetooth Low Energy Regulatory Aspects. (2011).

Page, S. R., Johnsgard, T. J., Albert, U., & Allen, C. D. (1996). User customization of a word processor. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 340–346). ACM.

Poynor, R. (1995). The hand that rocks the cradle: gillian crampton smith is making the royal college of art's computer related design program a multimedia powerhouse. *ID-NEW YORK-DESIGN PUBLICATIONS-*, *42*, 60–60.

re:form. (2014). Calm tech, then and now. Retrieved December 12, 2015, from http://www.johnseelybrown.com/calmtech.pdf

Rettig, M. (1994, April). Prototyping for tiny fingers. *Commun. ACM*, *37*(4), 21–27. doi:10.1145/175276.175288

Richardson, J. (2014). Double loop learning: a powerful force for organizational excellence.

Röcker, C. (2010). Social and technological concerns associated with the usage of ubiquitous computing technologies. *Issues in Information Systems*, *11*(1), 61–68.

Rogers, Y., Connelly, K., Tedesco, L., Hazlewood, W., Kurtz, A., Hall, R. E., … Toscos, T. (2007). *Why it's worth the hassle: the value of in-situ studies when designing ubicomp*. Springer.

Roos, T., Myllymäki, P., Tirri, H., Misikangas, P., & Sievänen, J. (2002). A probabilistic approach to wlan user location estimation. *International Journal of Wireless Information Networks*, *9*(3), 155–164.

Ruter. (2013). Evaluering av ny pris- og sonestruktur (nyps) i oslo og akershus. Retrieved from https://ruter.no/globalassets/dokumenter/ruterrapporter/2013/5-2013_evaluering_nyps.pdf

Ruter. (2014). Årsrapport 2014.

Ruter. (2015). Tilbud om alternative betalingsmidler for ruters billettsalg. Retrieved March 3, 2016, from https://ruter.no/globalassets/kollektivanbud/administrative-anbud/2015/rfp-alternative-betalingsmidler/rfp---betalingsmidler.pdf

Sakamura, K. (2003). Ubiquitous computing: making it a reality. *ITU TELECOM Worldfip. 1*, *9*.

Salas, A. C. (2014). *Indoor positioning system based on bluetooth low energy*.

Schade, A. (2015). Pilot testing: getting it right (before) the first time. Retrieved November 5, 2015, from https://www.nngroup.com/articles/pilot-testing/

Schilit, B. N. & Theimer, M. M. (1994). Disseminating active map information to mobile hosts. *Network, IEEE*, *8*(5), 22–32.

Shneiderman, B. & Plaisant, C. (2010). *Designing the user interface: strategies for effective human-computer interaction.* Addison-Wesley.

SIG, B. (2010). *Bluetooth specification version 4.0.*

SIG, B. (2014). *Bluetooth specification version 4.2.*

Skinner, E. A. (1996). A guide to constructs of control. *Journal of personality and social psychology*, *71*(3), 549.

Slettemeås, D. (2014). *Forbrukeres erfaringer medelektroniske billetter ogbillett-app'er til offentlig transport.* Retrieved March 20, 2016, from http://www.sifo.no/files/file79642_prosjektnotat_6-2014_web.pdf

SmartCardAlliance. (2014). Bluetooth low energy (ble) 101: a technology primer with example use cases. Retrieved from http://www.smartcardalliance.org/resources/pdf/BLE101-FINAL-053014.pdf

Sommerville, I. (2011). *Software engineering.* International Computer Science Series. Pearson.

Spiekermann, S. (2005). Perceived control: scales for privacy in ubiquitous computing. In *10th international conference on user modeling.*

Spiekermann, S. (2008). *User control in ubiquitous computing: design alternatives and user acceptance.* Shaker Aachen, Germany.

Stake, R. E. (2005). Qualitative case studies. In *The sage handbook of qualitative research (3rd ed.)* (pp. 443–466). Thousand Oaks, CA: SAGE.

Stevens, P. & Pooley, R. (2006). *Using uml: software engineering with objects and components.* Addison-Wesley object technology series. Addison-Wesley.

STMicroelectronics. (2015). *Nfc guide.* Retrieved from http://www.st.com/web/en/resource/technical/document/technical_note/DM00190233.pdf

Subhan, F., Hasbullah, H., Rozyyev, A., & Bakhsh, S. T. (2011). Indoor positioning in bluetooth networks using fingerprinting and lateration approach. In *Information science and applications (icisa), 2011 international conference on* (pp. 1–9). IEEE.

Svigals, J. (2012). The long life and imminent death of the mag-stripe card. Retrieved March 28, 2016, from http://spectrum.ieee.org/computing/hardware/the-long-life-and-imminent-death-of-the-magstripe-card

Tan, L. & Wang, N. (2010). Future internet: the internet of things. In *Advanced computer theory and engineering (icacte), 2010 3rd international conference on* (Vol. 5, pp. 376–380). IEEE.

Toftøy-Andersen, E. & Wold, J. G. (2011). *Praktisk brukertesting*. Cappelen Damm.

Torvmark, K. (2014). *Three flavors of bluetooth: which one to choose?* Texas Instruments.

van der Heijden, H. (2003). Ubiquitous computing, user control, and user performance: conceptual model and preliminary experimental design. *A Research Agenda for Emerging Electronic Markets*, 107–112.

Vlugt, E. (2013). *Bluetooth low energy, beacons and retail*. Verifone Headquarters.

Want, R. (2006). An introduction to rfid technology. *Pervasive Computing, IEEE*, *5*(1), 25–33.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *265*(3), 94–104.

Weiser, M. & Brown, J. S. (1996). Designing calm technology. *PowerGrid Journal*, *1*(1), 75–85.

Weiser, M., Gold, R., & Brown, J. S. (1999, December). The origins of ubiquitous computing research at parc in the late 1980s. *IBM Syst. J. 38*(4), 693–696. doi:10.1147/sj.384.0693

Appendix A

BLE packet structure

A BLE packet consists of several different fields (Lindh, 2015), as shown in Figure A.1.

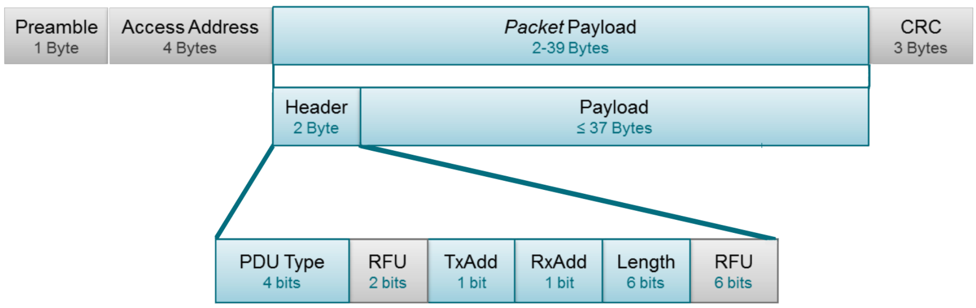

Figure A.1: Fields of the standard BLE packet (Lindh, 2015)

Preamble Used for synchronization between broadcaster and receiver.

Access Address: Address used to identify which communication link the packet is intended for with a 32-bit value that is randomly generated. For broadcast packets this field will always have the same fixed value showing that the packet is not meant for any one specific device.

Cyclic Redundancy Check (CRC): Used for error checking at the receiving end. Detects accidental errors on the packet e.g. caused by noise or interference.

Packet Payload: Consists of a header and a payload

The payload's header consists of several additional fields:

PDU Type: Determines the purpose of the device. For beacons this can be either connectable, non-connectable or non-connectable that provides additional information when responded to.

RFU: Blank fields set to zero which are Reserved for Future Use.

TxAdd: Single bit value indicating whether the address contained in the payload is public or private.

Length: The length of the payload.