NTNU
Norwegian University of
Science and Technology

Evaluation of the Accuracy of Positioning Data for Representing Room Usage

Tor Arne Lyngstad Holten

Master of Science in Computer Science
Submission date: June 2016
Supervisor: John Krogstie, IDI

Norwegian University of Science and Technology
Department of Computer and Information Science

Abstract

Norwegian University of Science and Technology (NTNU) is a large university consisting of several campuses, all containing many buildings and a variety of rooms. As a result of this magnitude it is challenging for facility managers to properly monitor and manage all building resources in the most advantageous approach. Since it's launch in 2011, MazeMap has helped both students and staff at NTNU to reach their destinations by providing means of outdoor and indoor navigation capabilities. In addition to being an application, MazeMap is also the underlying infrastructure for the collection of indoor positioning data from wireless devices at campus Gløshaugen. The creators of MazeMap, Wireless Trondheim, has also developed a system called Analytics System which uses the collected positioning data in order to calculate how many devices are located in each room at Gløshaugen. The goal of this research is to investigate whether or not the data produced by the Analytics System can provide a good representation of actual room usage at campus Gløshaugen.

The research uses the quantitative research approach in order to investigate if the accuracy of the collected positioning data is sufficient in order to give a good representation of room usage. By doing measurements in various rooms at Gløshaugen by measuring how many people are in the rooms it is possible to do comparison between actual room usage and the number of devices the Analytics System have estimated. Measurements are done in different types of rooms such as lecture halls, reading rooms and computer labs.

The result of the research is that the usage data produced by the current Analytics System can't be used to give a good representation of actual room usage at Gløshaugen. The numbers produced by the system is so low that no connection can be made between the calculated number of devices there are in a room and the actual number of persons in the room. It was only possible to see a small spikes in the usage data for measurements where about 150 persons were present in the room. Various approaches to improve the current Analytics System are suggested.

Sammendrag

Norges teknisk-naturvitenskapelige universitet i Trondheim (NTNU) er et stort universitet som består av flere campuser, og hvert campus består av mange bygninger som igjen består av mange forskjellige rom. Siden NTNU er så stort er det utfordrende for bygningsadministrasjonen å administrere alle bygningsressurser på den mest fordelaktige måten. Siden lanseringen i 2011 har MazeMap hjulpet både studenter og ansatte ved NTNU med å finne frem ved til forskjellige rom ved å tilby både utendørs og innendørs posisjoneringssystemer for navigering. I tillegg til å være en applikasjon er Mazemap også det underliggende infrastruktursystemet for innsamling av posisjonsdata fra trådløse enheter ved campus Gløshaugen. Skaperne av MazeMap, Trådløse Trondheim, har også utviklet et system kalt Analytics System som bruker de innsamlede posisjonsdataene til å kalkulere hvor mange forskjellige trådløse enheter det befinner seg i hvert rom på Gløshaugen. Målet med forskningen er å undersøke hvorvidt dataene som er produsert av Analytics-systemet kan brukes for å gi en god representasjon av den faktiske rombruken ved campus Gløshaugen.

Forskningen benytter seg av den kvantitative metodikken for å undersøke om de innsamlede posisjonsdataene er tilstrekkelig gode nok til å kunne gi en god representasjon som hvordan rom faktisk blir brukt. Gjennom å gjøre målinger i ulike rom ved Gløshaugen hvor det ble målt hvor mange personer som befinner seg i et rom er det mulig å sammenligne faktisk rombruk med de dataene som blir produsert av Analytics-systemet. Målinger har blitt utført i forskjellige typer rom som forelesningshaller, lesesaler og datasaler.

Resultatet fra denne forskningen er at dataene produsert av det nåværende Analytics-systemet ikke kan brukes til å gi en god representasjon av den faktiske rombruken ved Gløshaugen. Tallene som blir produsert av systemet er så lave at det er vanskelig å se noen forbindelser mellom hvor mange enheter det er beregnet å være i et rom og hvor mange personer som faktisk er i rommet. Ut i fra resultatene var det bare mulig å kunne se en liten økning i tallene fra Analytics-systemet når det var om lag 150 personer tilstede i et rom. Rapporten avslutter med forslag til ulike metoder som kan forbedre dagens Analytics-system.

Problem Description

Since its launch in 2011, MazeMap has helped students and staff at NTNU in navigating the large university. By using the existing Wi-Fi infrastructure, MazeMap is also able to collect positioning data from wireless devices located at or nearby the wireless network at Gløshaugen. The creator of MazeMap, Wireless Trondheim, has also created an system called the Analytics System which attempts to calculate how many devices are in each room at campus Gløshaugen based on the collected positioning data.

The task is to compare data produced by the Analytics System with how many people are actually using rooms at Gløshaugen in order to determine whether or not the Analytics System can produce usage data which can be used to give a good representation of room usage. The task is done according to the quantitative research methodology, and the report is expected to be written in English.

Preface

The work in this document is the result of research conducted in the Department of Computer and Information Science (IDI) at Norwegian University of Science and Technology (NTNU). The work is a part of the study program Computer Science and was carried out in the Spring of 2016.

I would mainly like to thank my supervisor John Krogstie at IDI for his guidance and contributions on this master thesis. I would also like to thank Dirk Ahlers at IDI for providing valuable feedback and ideas as a co-supervisor. Furthermore I would like to thank Håkon Lorentzen and Odd Erik Gundersen at Wireless Trondheim for allowing me to be able to get access to their Analytics System.

Trondheim, June 7, 2016

Tor Arne Lyngstad Holten

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AoA	Angle of Arrival
API	Application Program Interface
AP	Access Point
DGPS	Differential Global Positioning System
FM	Facility Management
GPS	Global Positioning System
IDI	Department of Computer and Information Science
IMU	Inertial Measurement Unit
IPS	Indoor Positioning System
JSON	JavaScript Object Notation
kNN	k-Nearest Neighbor
LAN	Local Area Network
MAC	Media Access Control
MSE	Mobility Services Engine
NTNU	Norwegian University of Science and Technology
OPS	Outdoor Positioning System
PDR	Pedestrian Dead Reckoning
PPMCC	Pearson Product-Moment Correlation Coefficient
RFID	Radio Frequency Identification
RF	Radio Frequency
RSSI	Received Signal Strength Indicator
RTT	Round-Trip Time

ToF	Time-of-Flight
URL	Uniform Resource Locator
WLAN	Wireless LAN

Chapter 1

Introduction

This chapter will give the reader a short overview of why and how the project was conducted. The chapter starts with describing background and motivation for the project. Following is the definition of the problem which defines the scopes of the project. Next, a description of the project is given. it gives an overview of the conduction of the project as well as the research outcome.

1.1 Background and Motivation

MazeMap has since it's launch in 2011 helped both students and staff at NTNU to reach their destinations by providing means of outdoor and indoor navigation. In addition to being an application, MazeMap does also provide the underlying infrastructure for the collection of indoor positioning data of wireless devices located at campus Gløshaugen. Position data is continuously being collected, stored and used for research purposes. There have been previous projects conducted at NTNU that have worked with the collected positioning data. Thingstad and Tran [2014] created a mobile service which helps students find available reading rooms at campus. Mersland [2013] created a prototype application for visualizing room usage, helping the administration to better understand how rooms are actually used. Eriksen [2015] and Aulie [2015], two projects conducted in parallel, investigated how visualization of the position data could aid facility managers, resulting in heat maps on top of maps to visualize human mobility.

Indoor positioning systems are used to locate objects position when outdoor positioning systems, such as GPS, are no longer able to accurately determine objects position indoors. MazeMap is able to determine an users indoor position with an accuracy of up to 5-10 meters [Biczok et al., 2014]. Despite this narrow range, some of the results from Aulie [2015] and Eriksen [2015] suggests there are certain areas at campus Gløshaugen where the accuracy of the position data is poor, especially in the entrance hall of the IT-building. The aim of this is to investigate the accuracy of the positioning data collected at NTNU Gløshaugen in order to get a better understanding of what factors affects the accuracy, as well as see how accurate the positioning data can be for representing actual room usage at Gløshaugen.

1.2 Problem Definition

The goal of this project is to determine if the collected positioning data can be used to give an accurate representation of room usage with the focus on campus Gløshuagen. This is achieved by collecting observational data of actual room usage on campus and comparing this data to the number of devices calculated by the Analytics System from Wireless Trondheim which does its calculations based on the continuously collected positioning data at Gløshaugen. The evaluation of the data should give insight into whether the accuracy of the position data does provide sufficient certainty, or if there are factors that cause the positioning data to be unreliable.

1.3 Project Description

This project was conducted over the period of 5 months in the spring of 2016. The work have been done in connection to work in Wireless Trondheim Living Lab [Andresen et al., 2007].

A preliminary project was conducted in the fall of 2015 as a pre-project to this master thesis. The focus was to perform a background study to identify previous work and research in the field of indoor positioning systems. A literature review was performed to uncover the current state of the art in the field, with a primary focus on research projects at NTNU which worked with the collected positioning data. The result from the preliminary project suggests that there is usefulness in using the collected positioning data for aiding facility managers with managing the vast room resources at NTNU.

The process of doing a background study have also continued in this project in order to further determine the focus of my research regarding indoor positioning data. Through meetings with supervisors it became clear that Wireless Trondheim was a potential stakeholder as they have developed their own system, Analytics System, for calculating how many devices are located in a room based on the collected positioning data. The focus of this project is to measure the accuracy of the collected positioning data by comparing the calculated device count from the Analytics System with actual room usage. By using the quantitative research method of observation, the researcher was able to conduct measurements in various rooms at Gløshaugen to gather data of actual room usage. Acquiring the data from the Analytics System were done through an API (Application Program Interface) which the researcher was able to get access to after a meeting with Wireless Trondheim.

The outcome of this project is to assess the accuracy of Wireless Trondheim's Analytics System to determine if the application can be used to give a good representation of actual room usage at campus Gløshaugen, and determining factors which affects the accuracy of the collected positioning data. The results of this project show that the Analytics System can't be used to give a good representation of room usage as the calculated device count is much lower than actual room usage. In some cases the number of devices estimated from the collected positioning data suggest that a room is empty, even though there are 40 or more people present in the room.

1.4 Report Outline

The outline of this report is:

- **Chapter 1 - Introduction** contains the background and motivation for the project, the problem definition and a project description.
- **Chapter 2 - Background** presents information of positioning systems, both outdoor and indoor as well as localization techniques. Then follows a description of Wireless Trondheim and their relevant applications such as MazeMap and the Analytics System. Lastly usages of indoor positioning systems are given.
- **Chapter 3 - Research Design and Methodology** described the chosen research methodology used over the course of the research. Both the goal and research questions are presented here.
- **Chapter 4 - Research Result** consists of the result from the research.
- **Chapter 5 - Discussion** is where the result is discussed, arguing whether or not the results can answer the given research questions given in Chapter 4.
- **Chapter 6 - Conclusion** concludes the thesis, and shortly describes the potential of future work.

Chapter 2

Background

This chapter contains the background study for the research. The chapter starts with a description of what positioning systems are. Various localization techniques, outdoor and indoor positioning systems are described in this section. The next section talks about Wireless Trondheim. The section covers the MazeMap system, how the position data is collected at Gløshaugen, and how the positioning data is processed by the Analytics System in order to generate usage data. Lastly, this chapter describes various usages of indoor indoor positioning systems. The section mentions various systems and research projects which utilizes indoor positioning systems. Research projects at NTNU that have previously worked with systems related to MazeMap is also presented.

2.1 Positioning Systems

A positioning system is a system used for determining the location of an object in space. There are two major categories of positioning systems: indoor positioning system (IPS) and outdoor positioning system (OPS). The application Mazemap (see section 2.2.1) utilizes both indoor and outdoor positioning systems when determining a users location. When indoors, the position is found by using the trilateration localization technique using wireless Access Points (see section 2.1.1.1) on campus. When the user is not in range of the wireless network the applation switches to GPS for finding the position.

2.1.1 Localization Techniques

A localization technique is a method for determining the position of an object. Positioning systems can employ such techniques individually or in combination. According to Hightower and Borriello [2001a] there are three principal techniques for automatic location sensing. These are *triangulation*, *scene analysis* and *proximity*. Fallah et al. [2013] does also describe the same techniques (but with different categories and names), including an additional technique called *dead-reckoning*[1]. The

---

[1] Estimates a users location based on last known position. While the user is moving, the dead-reckoning system estimates the users location through the aggregation of odometry readings from instruments such as accelerometers and gyroscopes. [Fallah et al., 2013]

same techniques are also described in [Gu et al., 2009], but with different categories for the techniques. As both Fallah et al. [2013] and Gu et al. [2009] focuses on indoor localization techniques, the techniques proposed by Hightower and Borriello [2001a] will be further described in this thesis as the techniques can be applied to both indoor and outdoor positioning systems.

2.1.1.1 Triangulation

Triangulation is a localization technique which uses the properties triangles to compute the location of an object. This technique can be divided into two subcategories. One is called *lateration* which uses distance measurements. The other is called *angulation*, resorting to the use of angle of arrival (AoA). *Trilateration* can be used to describe the method of using *lateration* for calculating an objects position, while *triangulation* refers to the *angulation* method.

Lateration

Lateration can calculate the position of an object by measuring the distance between the object and known reference points. In order to calculate an object's position in a three dimensional space it is necessary to have distance measurements from 4 non-coplanar reference points. Figure 2.1 shows how it is possible to calculate an object's position in a two dimensional space where only 3 reference points are needed. As the figure shows, a circle is drawn around each known reference point (*P1*, *P2* and *P3*). The radii of the circles (*r1*, *r2* and *r3*) is equal to the measured distance between the object at point B and to each reference point. The point where all circles intersects is the calculated position of the object.

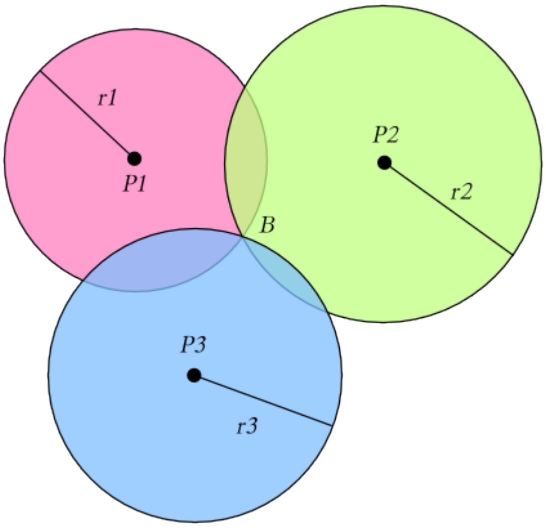


Figure 2.1: Each circle has a radius equaling to the distance between the object and the reference points. The location of the object is where all circles intersects.

Hightower and Borriello [2001a] describes three general approaches for measuring the distance required by the lateration technique. These are:

1. **Direct.** Measure the distance directly by using a physical action or movement. For instance, take measurements with a measuring band. It is challenging to automate direct measures as it requires coordinating autonomous physical movements.

2. **Time-of-Flight.** Time-of-flight (ToF) is used when measuring the distance between an object and a point P by using the time it takes to travel between the object and P at a known velocity. For instance, the speed of sound have constant speed given factors such as temperature, humidity and barometric pressure stays constant. Wong [1986] found the speed of sound in standard dry air at 0 °C and at barometric pressure of $101.325kPa$ to be $331.29ms^{-1}$. It is also possible to measure ToF of light and radio, but it requires clocks with very high resolution to get an accurate measurement. With the speed of light having a velocity of $299,792,458ms^{-1}$, the time it takes for light to travel 5 meters is about 16.7 nanoseconds.

   A challenge of using ToF to measure distances is that signals does not necessarily travel in a straight line between two points. Signals can be reflected, causing the distance traveled to be longer than the actual distance between objects, or refracted, meaning that the signal changes direction when traveling from one medium to another. Obstructions can also cause the signal to not reach its target.

   Another challenge is how known reference points agrees about the time. If an object sends out a signal, and the signal is reflected back to the object, the object now knows the round-trip time (RTT). The object in this case is both a sender and a receiver. In cases such as GPS (see section 2.1.2) the satellites are the sender and devices such as mobile phones are the receiver. Synchronizing the time becomes a major challenge when time resolution is high, and objects are continuously moving, changing the distance between objects.

3. **Attenuation**. When a signal travels, the intensity of it decreases as the distance from the signal source increases. Attenuation is the decrease in signal strength compared to the original intensity. The signal intensity of a radio signal in free space is attenuated by a factor proportional to $1/r^2$.

   Indoor radio propagation is difficult to predict because of the dense multipath environment and propagation effects such as reflection, diffraction, and scattering [Pahlavan, 2011]. A way to get around the radio propagation indoors is to use fingerprinting, a form of *scene analysis* (see section 2.1.1.2). Location fingerprinting associates location-dependent characteristics such as the received signal strength indicator (RSSI) with a location and uses these characteristics to infer the location [Kaemarungsi and Krishnamurthy, 2004].

Angulation

Angulation uses the angles between an object and reference points to calculate the objects location. It is necessary to know the distance between reference points to be able to calculate the position. In order to calculated the angle of arrival (AoA), it is needed to have multiple antennas, with known separation, measuring the time a signal arrives. As the distance between the origination of the signal to various reference points is different, a different ToA is measured for each reference point. Given when the signal arrived at the various reference points and the known distance between reference points, it is therefore possible to calculate the AoA to derive an objects location.

2.1.1.2 Scene Analysis

Scene analysis is a location sensing technique that uses features of a scene observed from a particular vantage point to draw conclusions about the location of the observer or of objects in the scene [Hightower and Borriello, 2001a]. Observed scenes are often simplified to contain features that are easy to represent and compare. There are two types of scene analysis: *static* and *differential*. In static scene analysis features of a scene is looked up in a data set where each feature has a known location. In differential scene analysis the difference between two scenes is followed to determine the position. When the observer moves, features will also move or change from scene to scene. Using the known location of features and how the features changed, it is possible to compute the position of the observer relative to the features.

Although scene analysis reminds of pictures with features on them, features can also be derived from other measurable physical phenomena. For instance, a feature can be the received signal strength indicator (RSSI) from an emitted radio signal. If the signal strength of each room in an office environment has a different signal strength it is possible to determine the position of a wireless device by looking up which room has what RSSI characteristics. This is often called a *fingerprint*, as each room has its own unique characteristic features.

2.1.1.3 Proximity

It is possible to derive whether an object is near a known location using the proximity location sensing technique. How near the object is depends in the approach to sensing the proximity. Hightower and Borriello [2001a] describes the following three approaches:

1. **Detecting physical contact.** This approach is a very low range approach. When an object is being sensed by a sensor with a known location, the location of the object is known. If the sensor doesn't sense the object, the object could be anywhere except at the sensors location. Examples of types of sensors for detecting physical contact includes pressure sensors, touch sensors and capacitive field detectors.

2. **Monitoring wireless cellular access points.** If a wireless device is connected to an access point, AP, the location of the device must be close to the AP. The area where the device could be depends on the signal strength of the AP. This approach to proximity sensing have a larger area where an object could be compared to detecting physical contact. The sensing area can

become very large in the instance of an cellular tower in a cellular network with a mobile phone connected to it.

3. **Observing automatic ID systems.** Systems such as a point-of-sale terminals for using credit cards can be used to find out where credit cards have been used if the point-of-sale terminals are in a known location. Where the credit card was before or after the transaction is unknown, but if used at another terminal the position can be inferred.

2.1.2 Outdoor Positioning Systems

An outdoor positioning system, such as Global Positioning System (GPS), is a system that provides geo-spatial positioning with a global coverage. In order to get coverage all around the earth, including all areas such as land, sea, air and space, it is needed to have satellite-based positioning for determining position [Hofmann-Wellenhof et al., 2007]. The satellites are artificially created and positioned in a way that there are always sufficient number of them that a position can be accurately determined anywhere on earth. For GPS, the minimum number of satellites that must be electronically visible is 4. This is achieved by having 21 evenly spaced satellites placed in circular 12-hour orbits inclined 55°to the equatorial plane [Hofmann-Wellenhof et al., 2012]. This ensures that there are at least 4 satellites visible anywhere on earth. Given the right location and conditions it is possible to have 10 satellites visible at a given time.

2.1.3 Indoor Positioning Systems

An indoor positioning system (ISP) can be used in areas where OPS no longer work optimally. The shortcomings of an OPS is in environments with obstructions, such as indoor areas. Walls, objects and even persons can obstruct the signals from an OPS. GPS is able to provide an approximation of an objects location when indoors if there are not many obstructions, but in the case of large buildings or underground tunnels, GPS signals could not reach devices inside. To counter the shortcomings of GPS, IPS can be used to provide a better mean for determining an objects location indoor.

There exists different types of IPS that work with various types of sensory information. Curran et al. [2011] mentions radio waves, magnetic fields and acoustic signals as possible ways to locate an objects position inside a building. Following are examples of each of the three mentioned types of sensory information:

- **Radio waves.** Radio Frequency Identification (RFID) utilizes radio waves for the tracking of objects. RFID tags can be active, meaning they have a power supply and regularly send out a signal, or passive, tags only emit a signal when activated by a scanning signal. Active tags can have a range of up to 100 meters while passive tags are often less than a meter [Curran et al., 2011].

- **Magnetic fields.** MotionStar DC, a magnetic tracking system from Ascension Technology Corporation, have a high accuracy and precision, on the order of less than 1mm spatial resolution. A disadvantage is that the magnetic sensors must be within 1 to 3 meters of the transmitter, and the accuracy further degrades with the presence of metallic objects in the environment [Hightower and Borriello, 2001b].

- **Acoustic signal.** AT&T Cambridge's Active Bats system uses ultrasound time-of-flight lateration technique to provide accurate physical localization [Hightower and Borriello, 2001b]. A short pulse of ultrasound is emitted from a transmitter which is attached to the object that is being located. The position of the object is calculated by measuring the time it takes for the signal to travel from the transmitter to receivers mounted at known locations in the ceiling. In order to get an accurate position of the object, it is required to have one ultrasound receiver to be installed every square meter [Addlesee et al., 2001]. The Active Bats system was deployed in a building, using 720 receivers to cover an area of $1000m^2$ on three floors. The accuracy of the system was around 3cm in three dimensions [Cambridge].

According to Liu et al. [2007] there are two approaches for implementing an IPS: One is to utilize an existing wireless network infrastructure, the other is to develop a signaling system and a network infrastructure of location measuring units for focusing on location an objects position wirelessly. The advantage of using an existing infrastructure is that it is cheaper and is less time consuming to set up. On the contrary, creating a custom infrastructure from the bottom up allows for control of physical specifications, therefore increasing the quality of the location sensing results. An example of a system that does not utilize existing infrastructure is the RFID system mentioned earlier in this section.

2.2 Wireless Trondheim

Wireless Trondheim[2] started out as a research and development project between NTNU, the city of Trondheim, the Sør-Trøndelag County Council, SpareBank 1 Midt-Norge, Adresseavisen and Trondheim Energiverk [Andresen et al., 2007]. One of the initiatives of the project was to create a Service Research and Development Lab. Wireless Trondheim and MazeMap has together with NTNU has set up a living lab by employing a Wi-Fi infrastructure in the city centre and NTNU's campuses [Andresen et al., 2007]. The living lab does not cover all the current campuses of NTNU as of the fusion of NTNU and the high schools in Ålesund, Gjøvik and Sør-Trondelag which happened 1 January 2016, but does instead cover the original NTNU campuses. The focus in this report is campus Gløshaugen.

2.2.1 MazeMap

MazeMap is developed by Wireless Trondheim and provides the underlying indoor positioning system at universities, hospitals, shopping malls and other large indoor venues. In addition to the infrastructure responsible for gathering users position,

---

[2] https://www.tradlosetrondheim.no/

there is also an application with the same name available for use on all types of devices: smart phones, tablets and computers. As an application, it is a wayfinding application that aids people in navigating the large campuses of NTNU to be able to reach their destinations. By being able to utilize GPS when the user is outdoors and an IPS when the user is indoors makes it a hybrid positioning application. When indoors, MazeMap can estimate a users indoor position with an accuracy of up to 5-10 meters [Biczok et al., 2014].

2.2.2 Cisco Mobility Services Engine

Cisco Mobility Services Engine (MSE) is the underlying service which provide location data to MazeMap. The MSE utilizes the existing network infrastructure at a location to provide location services. For campus Gløshaugen this is the "Eduroam" network. The Cisco MSE collects information about devices that are in range of the network. The discrete time, location and Media Access Control (MAC) address of the device is collected [Little and O'Brien, 2013]. When a device have Wi-Fi enabled it will continuously try to connect to the best AP nearby. This is done by periodically sending out a probe request frame. The Cisco MSE collect information about devices by reading the data in probe requests, which includes the MAC address of the device. When a device send a probe request, the request can be read by all APs nearby. The device is not required to be connected to the same network as the MSE for the MSE being able to gather probe request data. This also makes it possible to gather data from devices that is not connected to a Wi-Fi network, but has Wi-Fi enabled.

The frequency of probe requests sent from a device such as smart phones and tablets varies by device type, how long time it was since last request was sent, and whether the device is in idle or active mode. Idle mode is when the screen is off, while active is when the screen is turned on. Experiments conducted by Demir [2013] suggested that the frequency of probe requests are higher when a device is in active mode. The average interval of time between probes started out at 10s, but each cycle this value is multiplied by 2 until it reaches a fixed value. For idle mode the frequency is not regular. It can start out at every 10s, but after some time it could swiftly increase to 500s or 1000s.

The location of a device is calculated using trilateration (see section 2.1.1.1). More specifically it uses the method of attenuation: When each AP detects a probe request, information is collected about the content of the probe request and the RSSI, the signal strength of the received probe request. This information is sent to the WLAN Controller that manages all the access points, the further sent to the Cisco MSE which calculates the position of the device which sent the probe request. This flow of information is shown in figure 2.2.

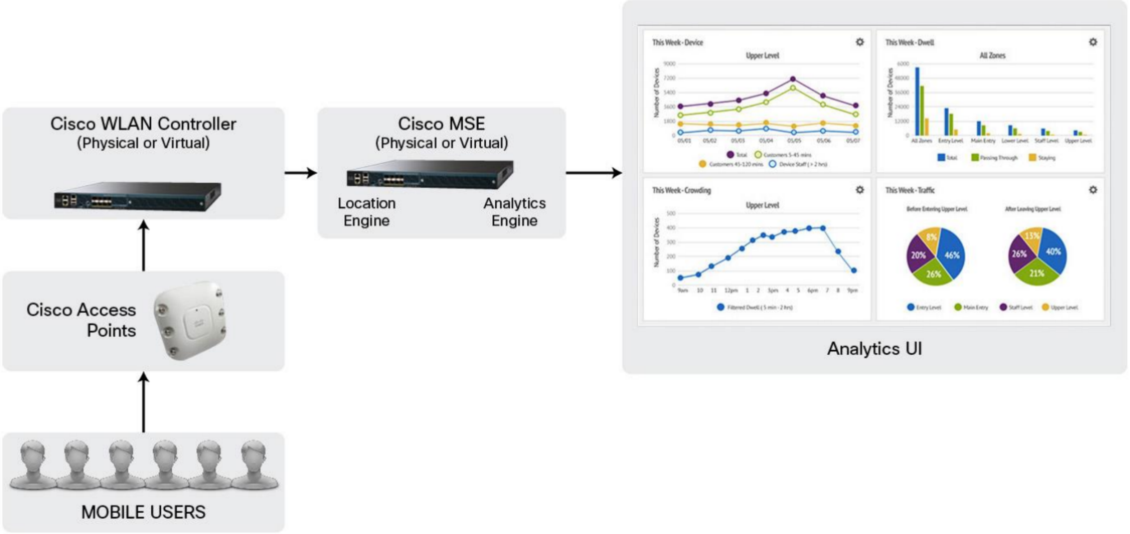


Figure 2.2: Overview of the Cisco MSE infrastructure. (`www.cisco.com`)

2.2.2.1 Privacy

The Norwegian Data Protection Authority[3] is a Norwegian Government agency, and is responsible for managing the Personal Data Act of 2000. The tasks of the agency is to ensure compliance with the Personal Data Act. Treatment of personal data, such as positioning data collected by the Cisco MSE, is strictly regulated. To be able to conduct lawful processing of positioning data there are, according to [UiO], three options:

- The service provider must have the consent of each person that positioning data is gathered for. How the data is going to be used must also be clearly be expressed for the persons being affected by the collection of data.
- The location-based service must be deemed necessary in order to fulfill an agreement. For instance: Insurance companies may require the installation of a GSM-system (Global System for Mobile Communications) when entering an insurance contract with owners of expensive cars.
- The service provider can process positioning data if they statutory authority to do so.

In the case of the Cisco MSE, none of the above options makes it lawful to collect and store the location data. In order to avoid these regulations and get the approval from The Norwegian Data Protection Authority, two products have been developed. The first product is called GeoPos. It is a service ensuring that a device using MazeMap is only able to get the position of the device MazeMap is running on, not the position of other nearby devices. MazeMap doesn't communicate directly to the MSE, but through the GeoPos that acts as a man-in-the-middle.

The second product is an anonymization module. In the event that somebody is able to get their hands on the collected positioning data, it is not possible to track

---

[3] `https://www.datatilsynet.no/`

a single user over a period of time. This is handled by not storing the original MAC address of a device. Instead, the MAC address is combined with a *salt* which is randomly generated every 24 hours. The combined MAC address and salt is then hashed, which creates a unique string which is stored instead of the device's MAC address. Since the salt only changes every 24 hours it is possible to track a single device for up to 24 hours, but it is impossible to continue tracking the same device the next day. This module has been approved by The Norwegian Data Protection Authority.

2.2.2.2 Cisco Hyperlocation Module

The Cisco Hyperlocation Module is an upgrade to existing Cisco Aironet Access Point. The module makes it possible for an AP to provide 1 to 3 metre location accuracy of associated Wi-Fi clients [Cisco, 2016]. At NTNU all APs are of a Cisco Aironet series, some of which is capable of the Hyperlocation module upgrade. Vidar Kværnø Stokke, operations manager of NTNUs network, has provided information regarding the various types of Cisco Aironet employed at NTNU. The current (2016.05.10) status of APs at NTNU (all of NTNU except the campuses in Ålesund, Gjøvik and Sør-Trøndelag) is shown in table 2.1. The *CleanAir* technology attempts to automatically identify source and location of interferences of the network, as well as taking action to avoid current and future interference, as well as full history reporting [Cisco].

Model Name	**Number of APs**	**CleanAir Capable**
All	1805	1802
AIR-CAP1702I-E-K9	2	2
AIR-CAP3502E-E-K9	113	113
AIR-CAP3502I-E-K9	819	819
AIR-CAP3702I-E-K9	871	868

Table 2.1: Overview of the current (2016.05.10) APs deployed at NTNU (all of NTNU except the campuses in Ålesund, Gjøvik and Sør-Trøndelag). Information is from Vidar Kværnø Stokke, operations manager of NTNUs network.

During the work on my thesis there had been certain areas at Gløshaugen where some Cisco APs has been upgraded with the Hyperlocation module. In May 2016, parts of the first floor of IT-bygget was upgraded. The lecture hall F1 with 6 APs has not been upgraded yet as the current APs in the room are mounted vertically on the wall. In order to get the best accuracy for determining a devices location it is necessary that the APs are mounted horizontally. The APs in F1 will be upgraded once a bracket mount solution to mount the APs horizontally on the wall has been made. Figure 2.3 shows how the Cisco Aironet AP without (left) and with (right) a Cisco Hyperlocation module installed.

Figure 2.3: Left is a Cisco Aironet AP 3700 series. Right is the same AP with the Hyperlocation module installed. (`www.cisco.com`)

2.2.3 Analytics System

The Analytics system is developed and maintained by Wireless Trondheim. The system is continuously calculating how many devices are in rooms at NTNU's campus Gløshaugen using the collected positioning data from the Cisco MSE (see section 2.2.2). The system is able to provide a device count at a fixed time interval, which is every minute. Through an Application Program Interface (API) refereed to as Analytics API, it is possible to obtain the usage data for almost all lecture halls, reading rooms, meeting rooms and computer labs at Gløshaugen. The algorithm that calculates the device count is property of Wireless Trondheim and cannot be disclosed in this report. But with permission the researcher is able to provide some details about how the algorithm works.

2.2.3.1 Filtering out Non-Stationary Devices

The algorithm does some filtering on what devices it considers when calculating the device count for each room. It only counts stationary devices, the ones that are not moving. As there are always some uncertainty with the accuracy of a device's location when finding its position using trilateration, it is possible that the algorithm counts a device as stationary if it has not moved a considerable distance from its previous location found by the Cisco MSE (see section 2.2.2). It is not known whether or not there are also some time limitations when determining if a device is stationary, for instance that a device must not have been moving in the last 5 minutes to be considered stationary.

2.2.3.2 Handling the Inaccuracy

The algorithm does also take a fuzzy approach when calculating device count for all rooms. When a device sends out a probe request, the Cisco MSE calculates the position of a device. In addition to the position, an accuracy value is also calculated. This value says something about how certain it is that a device is in the location which was found using the trilateration localization method. When the Analytics System have filtered out all non-stationary devices, it starts working on calculating how many devices there are in each room. The fuzzy approach for this is that each device has as mass value. This mass is distributed among the room which the device is located in according to where Cisco MSE thinks it is, and the nearby rooms. If the accuracy value related to a device's position suggest a low accuracy, more of the mass is distributed among the nearby rooms. For instance: a device has a total mass value of 1.0, and it had a low accuracy value. This could means that only 0.6 of this mass is added to the room corresponding to the position of the device, while the remaining 0.4 is distributed evenly among the nearby rooms. Every device has a total mass of 1.0 it can add to rooms, and the algorithm considers nearby rooms as the adjacent rooms. When the algorithm have iterated over all devices, each room will now contain a mass value. This mass value is now the same as the number of devices the system think is in said room.

2.2.3.3 Interfering Processes

About twice a day there is various processes that is running on the collected positioning data, such as backup of positioning data. The Analytics System is not able to calculate device count when these processes are running. It it not fixed exactly when these processes start and stops, but they usually run at midnight and at noon, give or take a few hours. The Analytics System is still running when these processes run, and it is possible to perform queries to the Analytics API to get data, but the response is that the usage is calculated to be 0 usage instead of something descriptive as *No Data* when the other processes had been going on. During the project the researcher has noticed that the processes which run at around midnight and noon usually take about 30 to 40 minutes to complete.

2.3 Usage of Indoor Positioning Systems

This section contains various examples of usages of indoor positioning systems (IPS). The section presents some commercial usages, examples of research projects where IPS has been the focus, and related research projects at NTNU.

2.3.1 Commercial Use

This section describes some systems that utilize indoor positioning in commercial products. Although MazeMap could be put in this section as it is both a research project and useful for people to navigate various venues around Norway, it has already been described thoroughly in section 2.2.1.

2.3.1.1 Google Indoor Maps

Indoor Maps is developed by Google. It works similarly to Google Maps which allows users to look upon maps and satellite images from all over the world. The difference is that Indoor Maps makes it possible to view indoor floor plans for various venues. Indoor Maps is not a stand alone application, but is integrated into Google Maps. By zooming in on buildings and venues, the floor plans becomes available, although it is only possible if floor plans have been added to buildings. It is possible for anyone to add floor plans through the Google Indoor Maps web page [Google]. Figure 2.4 is an example of Madison Square Garden, where on the left is how the venue appears on Google Maps, while on the phone to the right is how it looks when viewing the same area using Google Indoor Maps.

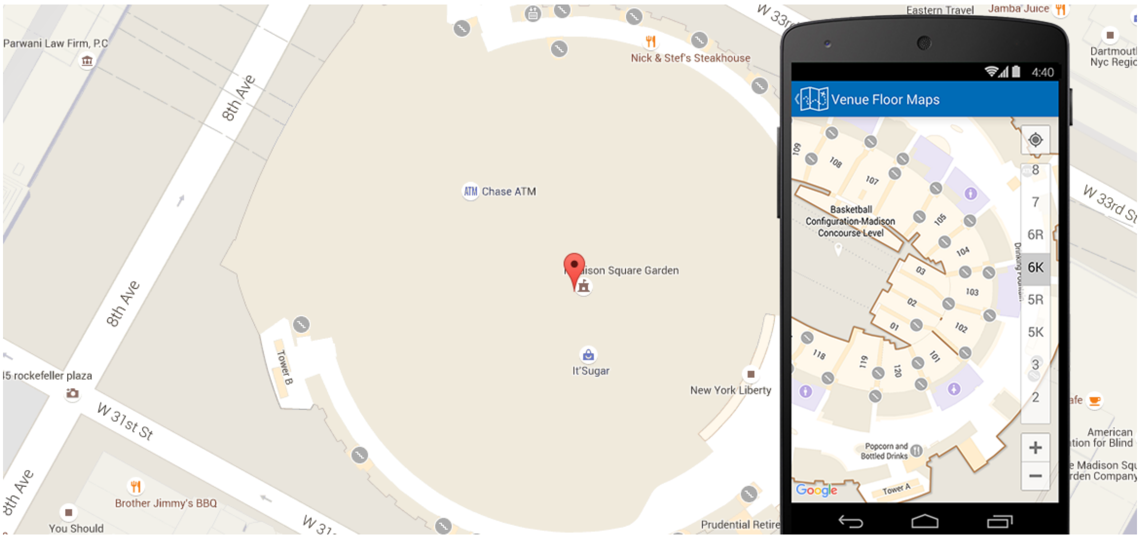


Figure 2.4: Indoor Maps makes it possible to view and navigate floor plans of buildings and venues, such as Madison Square Garden. (`www.google.com`)

2.3.1.2 Wifarer

Wifarer is a company that specializes in the development of mobile applications and cloud solutions for businesses by utilizing indoor positioning algorithms and navigation technologies. Wifarer has a number of clients, such as hospitals, airports, museums, stadiums, shopping centers and universities around the world [Wifarer]. Clients usually have a mobile application for helping their customers navigate their often large venues. Wifarer have a content management system which their clients can use to customize their app's look and feel, add directories, and attach rich media to any indoor location. The latter of providing additional media at a location can be really helpful in museums if visitors want to read about more detailed information of paintings and exhibitions.

2.3.2 Research Projects

This section contains research projects where IPSs have been used. The goal is not to do an extensive literature review of the field, but present a general overview of relevant research projects where IPS and location data has been a main part of the project.

2.3.2.1 IPS using Accelerometry and Heading Sensors

A project by Collin et al. [2003] explains that the majority of previously proposed personal positioning systems utilize an approach called Pedestrian Dead Reckoning

(PDR). This approach tracks a person by using accelerometers for step detection and step length estimation, and magnetic compasses or gyros for determining the direction the person is heading. They proposed a system which used a tactical-grade IMU (Inertial Measurement Unit) for keeping track of heading information and accelerometers for step occurrence detection. It was tested if the number of gyros would affect the accuracy using PDR.

A dead reckoning system requires that an object has a known location prior to using only accelerometers or gyros for keeping track of the change in position. In the research, Collin et al. [2003] found that using differential global positioning system (DGPS) and three sensors to give the best accuracy. If the DGPS was not available, or using only a single gyro, the accuracy was poor.

The outcome of the work was that the system they proposed would be limited due to size restrictions, but they think that it will be more applicable in the future as gyro technology continues to evolve, making the required hardware small enough. As the project was from 2003, the technological progress since then has made it possible to fit gyros inside devices such as smart phones. Potential usages for the system was to find the position of rescue workers, police squads, and other indoor location and navigation applications.

2.3.2.2 Analysis on IPS Technologies for Supporting Facility Management

Taneja et al. [2010] did a research project on the analysis of three indoor localization techniques for supporting facility management (FM) field activities. The three localization techniques tested was Radio Frequency Identification Tags (RFID), Wireless LAN (WLAN), and Inertial Measurements Units (IMUs). These three techniques was selected as they did not requires line-of-sight, was scalable and low cost. The authors used a fingerprinting approach, which means the creation of a signal strength map for localization (see section on scene analysis 2.1.1.2). The k-Nearest Neighbor (kNN) algorithm was used for location determination. Accuracy and precision was tested and measured for each of the three localization techniques. The testing was performed on the same floor of an actively used university building. The test for IMU was performed following a specific route in the same location.

The results from the paper suggests that WLAN satisfied all the requirements of indoor localization technology identified in the paper. Table 2.2 shows the accuracy and precision found during the testing. Another thing to note is that when the WLAN test was repeated with only 9 different access points instead of the original 32, the accuracy fell only to 3.05m for 95% confidence. The poor results for RFID is not known, but it is suspected it could be a result of the RFID tags used for the test was being read only at close range.

2.3.2.3 COMPASS: Accelerometry and High Accuracy Heading Sensors

King et al. [2006] developed a positioning system called COMPASS which is based on 802.11-compliant network infrastructure and digital compasses. On a mobile device, COMPASS samples the signal strength values of the access points within range of the device and uses the compass to determine the orientation of the user.

Technology	**Average accuracy of 5 samples**	**95 percentile accuracy**
WLAN	0.58m (Best case kNN)	1.52m (Best case kNN)
RFID	11.9m (Best case kNN)	>30m
IMU	13.1m drift error at the end of 150m route	18.6m

Table 2.2: Table 2 from [Taneja et al., 2010] shows the result of testing three different localization technologies when determining an objects location.

An probabilistic positioning algorithm can learn the various fingerprints of signal strength (see section on *scene analysis* 2.1.1.2) which varies depending on the orientation of the person. This is because the person is an obstruction for the wireless signal. After a short period of training, the COMPASS system had an average error distance of less that 1.65 meters in an experimental environment of 312 square meters. COMPASS shows that combining wireless signal fingerprinting with an objects orientation can help on mitigating the obstruction factor which affects positioning systems which uses radio signals for location determination.

2.3.2.4 DOLPHIN: An Autonomous Indoor Positioning System

DOLPHIN was developed by Fukuju et al. [2003] in order to deal with the problem of providing accurate physical location tracking in large-scale spaces for ultrasonic sensor nodes. The problem was that it required a lot of manual configuration for all sensor nodes in order to provide sufficient accuracy when determining an objects position. The DOLPHIN system consists of distributed wireless sensor nodes which is capable of sending and receiving both radio frequency (RF) and ultrasonic signals. The RF function is used for synchronizing time and exchanging of messages between nodes, while ultrasonic signals is used for measuring the distance between nodes.

A simple explanation of how the system works: When node A transmits an RF signal containing a message that other nodes know it is node A that sent the signal, all nearby nodes starts an internal pulse counter. At the same time as sending the RF signal, node A does also send an ultrasonic signal. When the sound signal from node A reaches another node, the receiving node notes down how long it took for the signal to travel from node A to itself, thus enabling it to calculate the distance to node A since the speed of sound and time taken is known. If the same node is also able to calculate the distance to two other nodes, it is possible to calculate the position of the node using trilateration (see section 2.1.1.1).

In order to calculate the positions of nodes in a large network of nodes, it is required that some nodes have a known location. These are called reference nodes. The autonomous part of DOLPHIN is that all nodes which are not reference nodes is able to automatically find their own position in the world based on the reference nodes, meaning that there is no need for manual configuration of each node. However, this requires that there is a sufficient number of nodes distributed such that it is always possible for a single node to be able to receive both RF and ultrasonic signals from at least three other nodes. If all nodes are located in the same plane it is only required to have two nearby nodes when calculating a nodes location.

The researchers evaluated the performance of the system to be able to find a position with an accuracy of around 15cm. However, the accuracy did degrade if the hop count was large. The hop count is how many nodes there are between a node and a reference node. A small error in position localization occurs every time a node calculates its position. Error propagation occurs when the position of a node is based on another nodes position which already has a small error in position. The more hops there are between a node and a reference node the larger the error in position becomes. The reason for errors in localization is due to factors such as obstruction, reflection and attenuation of signals.

2.3.3 Projects at NTNU

This section describes research projects that have been previously conducted at NTNU. All projects are related to IPS.

2.3.3.1 Ledig Lesesal Plass

A master thesis by Thingstad and Tran [2014] created a mobile service that uses anonymous indoor location-based data. The name of the application is *Ledig Lesesal Plass* (*Available Reading Room* in English). The purpose of the application was to help students finding available reading rooms, something which is especially hard during the exam periods when most reading halls are filled up with students preparing for exams. Figure 2.5 shows a mockup of the application they developed. The figure shows a list of available reading rooms, an estimate of how full each room is and the distance from where the user currently is located and to the room. The room usage estimation data is from the Analytics-API (see section 2.2.3), while the users current position is found using MazeMap's system for finding a device's location. The room usage estimation is based on how many wireless devices there are in the room, such as smart phones, tablets and laptops.

2.3.3.2 Visualizing Room Usage

A prototype application by Mersland [2013] was developed in the spring of 2013. The purpose of the application was to visualize how rooms are utilized on NTNU's campus Gløshuagen. The application used location data from the Analytics-API (see section 2.2.3) to do this. The goal of the thesis was to see if such an application could help administration to get a better understanding of how rooms are actually used on campus. Rooms can be reserved[4], and it is therefore logical to think the room have been used during the time it was reserved. The application by Mersland could help by allowing facility managers to know if reserved rooms actually had been in use instead of being left empty for the duration of the reservation. Figure 2.6 shows a graph in the application which is still running today on one of Wireless Trondheim's servers. The screenshot was taken 2016.03.12 at about 14:45.

---

[4] `https://romres.ntnu.no/`

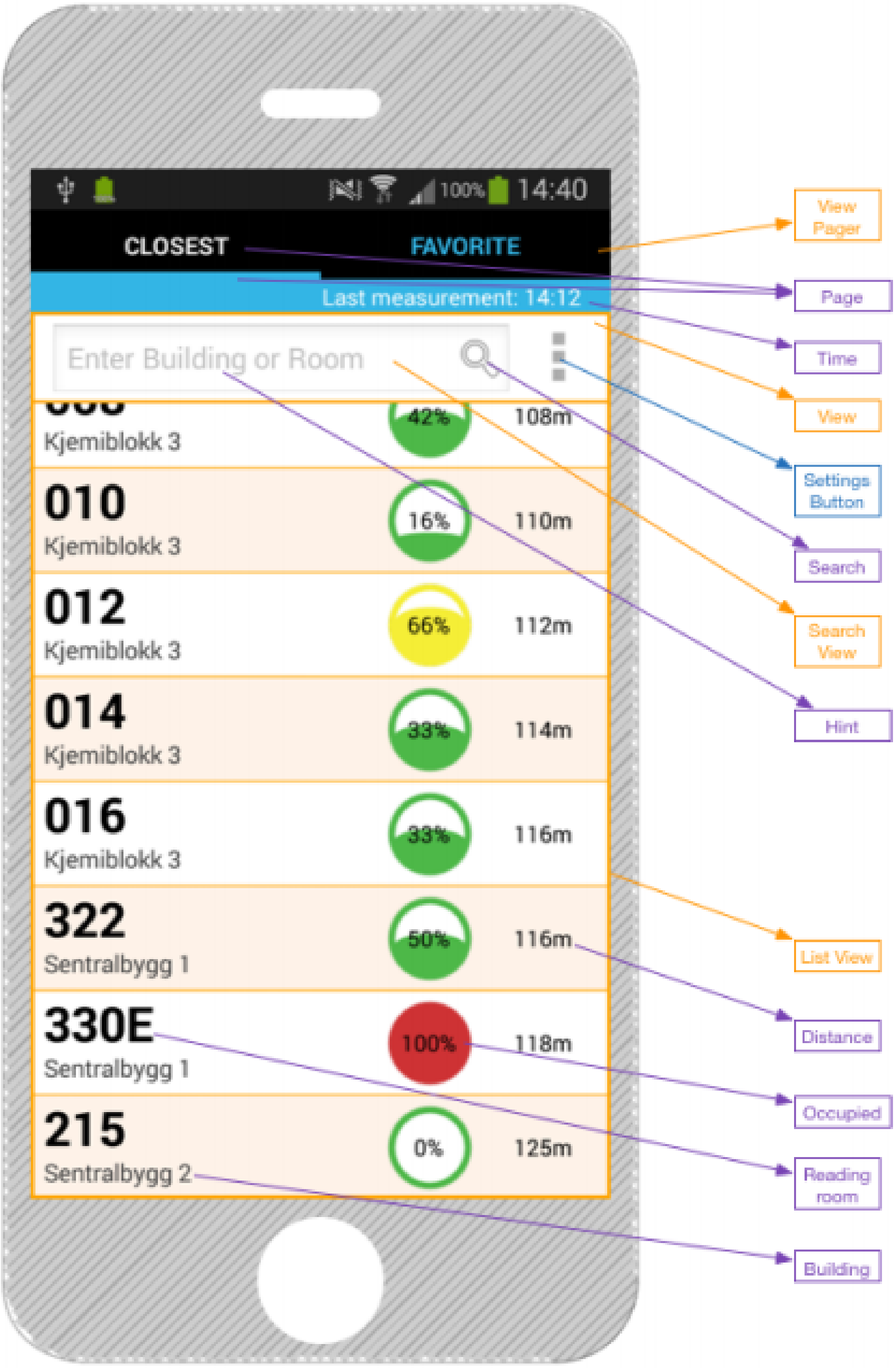
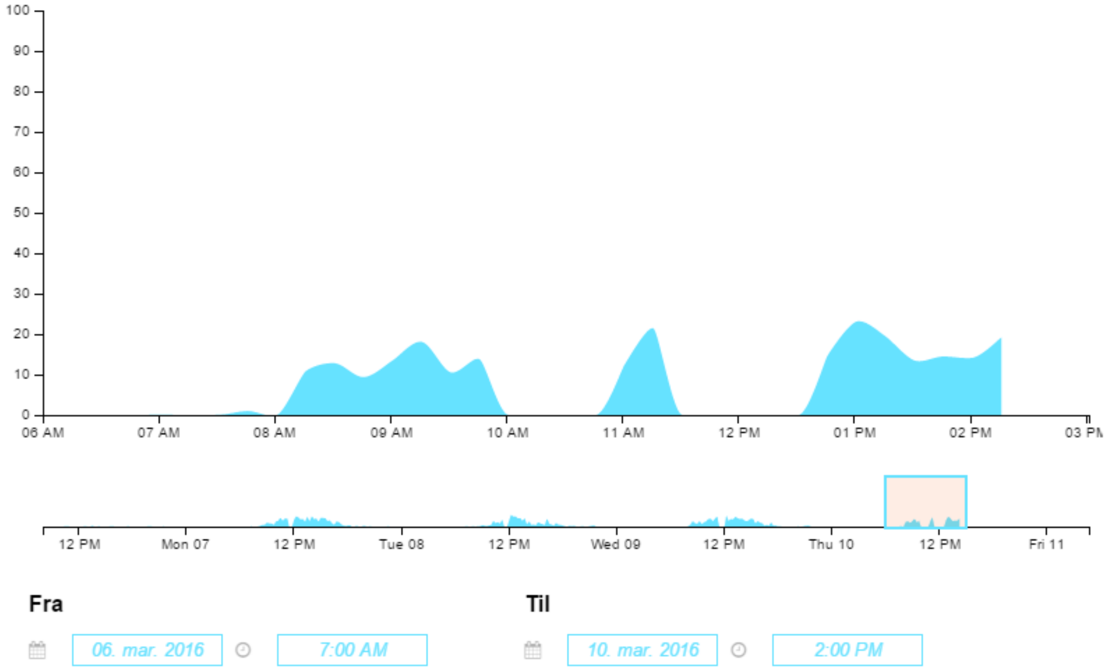


Figure 2.5: A mockup of the main screen of Ledig Lesesal Plass showing a list of reading rooms. An estimation on how full the rooms are is shown for each room. [Thingstad and Tran, 2014, Fig. 6.2]



Fra
06. mar. 2016 7:00 AM

Til
10. mar. 2016 2:00 PM

Figure 2.6: A screenshot from the prototype developed by Mersland [2013]. The graph shows the number of devices the Analytics-system has calculated to be present for the room S5 (in Sentralbygget) on 2016.03.12 from 06:00 up to 14:45. It is possible to look at historical data using the application.

2.3.3.3 Visualizing Human Mobility Patterns

Two master theses were conducted in the spring of 2015. These are by Eriksen [2015] and Aulie [2015]. Both investigated how positioning data could be used to aid facility managers. This was achieved by developing prototype applications which visualized the positioning data on maps as heat maps to show where people are located at various times. Being able to choose position data from specific time intervals it is possible to investigate human mobility patterns at different times of the day, or at different days of the week. In both theses the conclusion was that the information form the prototypes is useful for stakeholders such as facility managers. The positioning data used is the data collected from the Cisco MSE (see section 2.2.2).

Figure 3.1 and 3.2 in section 3.3.1 shows excerpts from the prototype developed by Aulie [2015].

2.3.3.4 Mazemap and Wayfinding Requests

Martinez [2013] did analyze the wayfinding requests made in the MazeMap application in order to find users' mobility patterns, traffic bottlenecks and popular paths taken on campus Gløshaugen. The result was visualized using graphs showing relationships such as what building is most often the goal when a person is in a specific building, and what buildings are used most. The conclusion was that most of the outcomes could be predicted by looking at the campus distribution, but the result of the project verified this.

Chapter 3

Research Design and Methodology

The contents of this chapter is the chosen research design and methodology of the conducted research. The chapter starts off with presenting the goal and research questions which leads the research. Then follows the reasoning of using the quantitative research approach, as well as how the research process was conducted. The chapter does also describe how data is collected and what rooms were selected for doing measurements in. Finally, methods for comparing the collected data is presented.

3.1 Goal and Research Questions

The aim of the Thesis is to investigate if the collected positioning data can be used to give an accurate representation of room usage at campus Gløshaugen. This is achieved by collecting observational data of actual room usage on campus and comparing this to the number of devices calculated by the Analytics System from Wireless Trondheim. This leads to the following research goal:

> **Goal**: Investigate if the number of devices calculated from collected positioning data can give an accurate representation of room usage in rooms at campus Gløshaugen.

With this goal in mind, the following research questions are derived. There are two research questions, RQ1 and RQ2. The first question does also have two sub questions that are related to RQ1. These are RQ1.1 and RQ1.2. The first research questions inspects whether the calculated number of devices for a room can be used to derive a good representation of actual room usage. The second research questions focuses on finding how the accuracy of the positioning data can be improved.

> **RQ1**: How good representation of actual room usage does the calculated device count from the Analytics System provide?

This research question seeks to answer the goal of this research. If the collected positioning data are accurate, then the calculated device count will give a more correct measure of how many devices are located in a room. If there are devices present in a room it suggests that there are people using it.

> **RQ1.1**: How does the representation of room usage change depending on the type of room being used?

This is a sub question to RQ1 that further explores if the room type affects the accuracy of calculated device usage representing actual room usage. Different room types could mean people use the rooms differently. It is reasonable to believe that a student used a various number of personal devices depending on the purpose of a room. The Analytics API can provide room data from four different categories. These are lecture halls, reading rooms, meeting rooms and computer labs.

> **RQ1.2**: What is the linear relationship between the calculated device count and the number of people in a room?

Another sub question to RQ1 aims to discover if there exists any linear relationship between calculated device count and the number of people in a room. Is it possible to find the average number of devices used per occupant of a room? It is also possible to explore if the average changes depending on room type.

Finding a linear relationship can make it possible to determine actual room usage from the calculated number of devices in a room. For instance, if the number of devices in a room is 10 while the average device per occupant is 0.2, the number of people in the room is expected to be 50.

> **RQ2**: How can the accuracy be improved or the inaccuracy be better handled?

Finding and determining factors that affects the accuracy of the position data can help in improving the accuracy in rooms and areas where the accuracy is poor. Instead of improving the accuracy, discovering ways of handling the inaccuracy could make make the calculated device count give a better representation of actual room usage.

3.2 Research Approach

The research approach chosen to answer the research questions is using the quantitative research approach. This section gives insight into what quantitative research is about, following with a brief description of the data collected through this research.

3.2.1 Quantitative Research

The quantitative research approach focuses on using a large amount of data to answer the research questions. Quantitative data is described by Hox and Boeije [2005] as "Data that can be described numerically in terms of objects, variables, and their values". The data is often gathered by structured means from a population, which makes the collected data a sample of the population. It is often impossible to gather data from a whole population due to its size. A sample becomes a better representation of the population the larger it is. As the data collected in quantitative research is numerical, it is possible to perform analysis using mathematically based methods [Muijs, 2010]. By doing analysis on quantitative data it is possible to derive relationships between one thing and another within a population.

The data which is collected throughout this research project are all related to room usage on campus Gløshaugen. There are primarily gathered two types of data: How many people that actually uses a room and how many devices Wireless Trondheim's Analytics System (see section 2.2.3) thinks there are in a room based on positioning data.

Number of people
This is the number of people present in a room at a given time. This includes all occupants such as students, teachers, janitors and the researcher if the data is gathered from within the room. Obtaining this data requires the researcher to observe a room and count how many occupants there are.

Calculated device count
This is a number which is generated by Wireless Trondheim's Analytics System. Every minute the system calculates the device count for every room on campus Gløshuagen based on the indoor positioning data gathered continuously from probe requests sent from devices. See section 2.2.2 for more information of how position data is anonymously collected, and section 2.2.3 on a more detailed description of how the Analytics System works.

3.3 Research Process

This section describes the research process for the project. The goal is to present an overview of the execution of the project.

3.3.1 Background and Motivation

A preliminary research project was conducted in fall 2015, the semester before this master project. The aim of the pre-project was to do a literature review related to IPSs in order to determine what the researcher would work on during the master project. The result from the literature review was an idea of creating an artifact which used the positioning data in order to show how many people are present in a room. The artifact would present a map of campus, and depending on how many devices are in various rooms, the rooms can be colored according to how full a room is. For instance green would represent a empty room, and red would mean the room is at full capacity. A gradient color between green and red could be used to indicate that a room is at half capacity.

This type of application could help facility managers with the managing of the many hundreds of rooms at NTNU Gløshaugen. During the start up phase of this master thesis it was found out that creating such an artifact would be very similar to a previous project. Thingstad and Tran [2014] developed a prototype mobile application which helped students find available reading rooms. There are more students than reading spots at Gløshuagen, and this makes it difficult for student to find a spot, especially during exam periods. The application used the data produced by Wireless Trondheim's Analytics System in order to determine how many people are in a room.

As the original idea would be too similar to a previous project [Thingstad and Tran, 2014], the researcher became more interesting in assessing the accuracy of the data produced by the Analytics System. This would also assess the accuracy of the positioning data which is collected by the Cisco MSE (see section 2.2.2). Previous work done by Eriksen [2015] and Aulie [2015] did suggest that there was areas at Gløshaugen where the accuracy of the positioning data was poor. Figure 3.1 shows an excerpt from the application developed by Aulie [2015]. The positioning data is visualized using heat maps. The large cluster in IT-bygget is most likely the people located in the large lecture hall F1, but due to poor accuracy of the positioning data the mass of people is believed by the Cisco MSE to be located in the entrance hall and other room instead of the F1 where it is most likely where people are located. The visualization shows all registered positions from 10:00 to 14:00 the 1st of October, 2014. Eriksen [2015] did also find that the accuracy was poor in the same location.

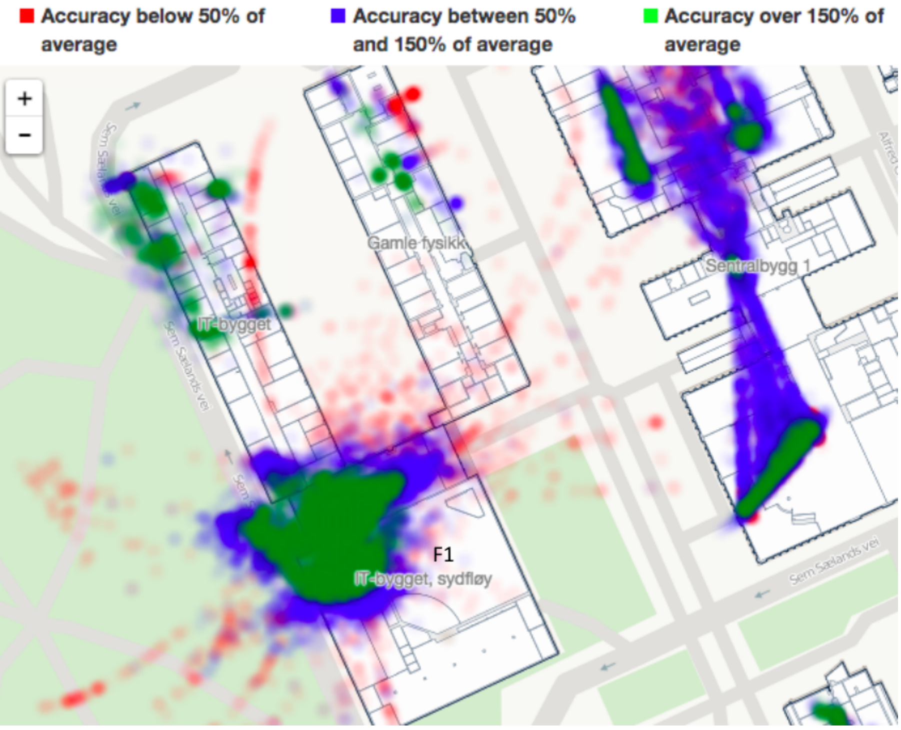


Figure 3.1: An excerpt from the application developed by Aulie [2015, Fig 3.10, p. 37], displaying the accuracy of all registered positions from 10:00 to 14:00 the 1st of October, 2014.

Another figure from Aulie [2015], figure 3.2 in this report, did also show another interesting location regarding the accuracy of the positioning data. Figure 3.2 shows a similar heat map, but this time for the 13th floor of Sentralbygg 1 (bottom) and

Sentralbygg 2 (top). There is a clear difference of the accuracy of the positioniong data between the buildings. For Sentralbygg 2 all the positions are located within the walls of the building. The same thing can't be said of Sentralbygg 1 where some of the positioning data is located outside of the actual building. Since the figure shows the 13th floor it is certainly that some of the positioning data is wrong as there can't be people outside the building at such a height. If it had been the ground floor it would be possible as the Wi-Fi-signal can reach persons outdoors. A major motivation for this project has been to investigate what factors causes the accuracy of the positioning data to vary so much around at Gløshaugen.

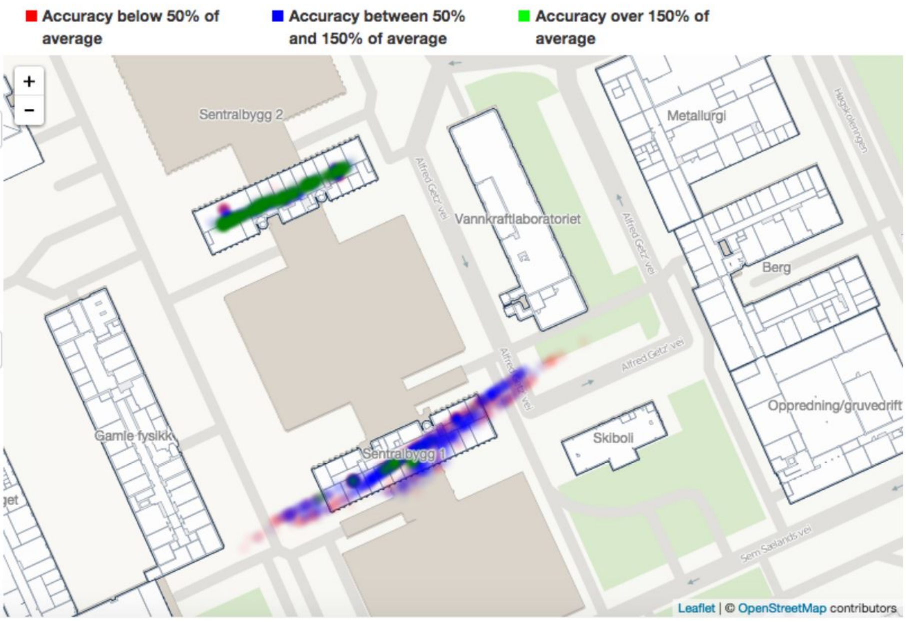


Figure 3.2: The accuracy heatmap of the 13th floor for Sentralbygg 1 and Sentralbygg 2. [Aulie, 2015, Fig 5.10, p.67]

As the Analytics System calculates how many devices are in each room based on the same positioning data which Aulie and Eriksen visualized, it is possible to measure the accuracy of the positioning data by comparing the device count produced by the Analytics System with real world data of how many people are located in a room.

3.3.2 Analytics System

When the scope of the project had been narrowed down to investigating the accuracy of the positioning data, the researcher and co-supervisor Dirk Ahlers planned a meeting with Wireless Trondheim about getting to know more about the Analytics System and hopefully get access to it. The meeting consisted of the researcher, Dirk Ahlers and Håkon Lorentzen, whereas the latter became the researcher main person of contact for further communication regarding Wireless Trondheim. The outcome of the meeting was that the researcher got access to the Analytics API, and that another meeting was planned where the researcher were to visit Trådløse Trondheim and being allowed to look at the source code for the Analytics System in order to learn more about the algorithm which calculates device count for all the rooms at Gløshaugen. Notes were taken during this meeting, and the details of the system is presented in section 2.2.3.

After the researcher had gotten access to the system, time was spent in order to learn about how data could be extracted. There was no official documentation of how to use the Analytics API, but the researcher were able to look at the previous projects, such as Thingstad and Tran [2014] and Mersland [2013], to see how data could be extracted from the system. Obtaining data was done by changing an Uniform Resource Locator (URL) and its parameters in order to specify details such as which room to get data from, and time interval for the data. In addition, a *points* parameter could be set in order to specify how many data points would be returned from the specified time interval. More information of how the URL looks like and what options the researcher had when extracting data is presented in section 3.4.2.

3.3.3 Measurements

After the researcher had learned about how the Analytics API worked and how to extract data for specific rooms and time intervals, work was started on doing measurements in rooms at Gløshaugen. Fiol was measured first as this room was where the researcher primarily spent working on the project. Fiol was also the room which was measured more than once as it was convenient to do so for the researcher. After the researcher had successfully conducted measurements in Fiol and managed to compare the calculated device count with the number of people present in the room, the researcher started work on doing measurements in other rooms at Gløshuagen. Lecture halls was prioritized as it was desirable to conduct measurements when lecture halls were still in use for lectures. Reading rooms and computer labs were deprioritised as these rooms would still be in use after lectures stopped, which is often sometime between week 15 and 17 for most subject.

When a measurement had been conducted, the researcher had data about actual room usage. The next step is then to get matching data from the Analytics API and collect both data in a joint table in Excel which makes it possible to create a graph to visualize any correlation between the data. Calculating the Pearson Product-Moment Correlation Coefficient (see section 3.5) for the data sets were done by using Excel's built in function called *correl*.

3.4 Data Collection

In the research there are two primary sources of data which is gathered. The first is how many people there are in a room, and the second is how many devices the Analytics System have found to be in the same room based on the collected positioning data from the Cisco MSE (see section 2.2.2). This section describes how data was gathered as well as which rooms was chosen for the measurements.

3.4.1 Number of People

This data is how many people are present in a room at a given time. In order to obtain this data the researcher decided to do measurements in rooms by being present in the room in order to being able to observe the people in the room. This limits the researcher to only measure rooms which the researcher has access to, and rooms where the researcher will not disturb the ongoing activity.

3.4.1.1 Room Types

There are four room types available in the Analytics System. These are *lecture halls*, *reading rooms*, *computer labs* and *meeting rooms*. In order for the researcher to do measurements about how many people are in a room it is necessary that the researcher is present in the room to be able to get an accurate count of the number of people inside. The researcher must also have access to the room in order to do a measurement in it. The chosen approach by the researcher is to be anonymous while doing the measurements. Lecture halls are a great example of rooms where the researcher can conduct measurements while being anonymous as it will appear that the researcher is one of many students taking part in the lecture. It is also possible for the researcher to join a lecture even if the researcher is not enrolled in the subject the lecture is about. This makes it possible for the researcher to conduct measurements in lecture halls all over Gløshuagen. On the other hand, meeting rooms can not be booked by students, and the activities in these rooms makes it hard for me to do measurements anonymous. Therefore the researcher have decided to not do measurements in meeting rooms.

The remaining room types, reading rooms and computer labs, are more restricted for the researcher to access. Some reading rooms and computer labs can only be accessed by students from certain departments. Every student have an access card that can be used to unlock doors at NTNU. Cards are programmed with various permissions, meaning that only specific rooms and areas can be unlocked by a specific access card. The permissions depends on what department a student is in, but it is possible to add permissions in some cases. For instance: When applying for, and being granted, a seat in one of the reading halls reserved for master students at the Department of Computer and Information Science (IDI), the access card is updated with permissions making it possible to access these rooms.

The reading rooms and computer labs the researcher have access to are the ones that are available to all students, and the ones that are reserved for master students at IDI. In Realfagsbygget, the largest building on Gløshaugen, all students can enter and leave the rooms B1-101, B1-105 and B1-165 without needing to use access cards. There are several other reading rooms in Realfagsbygget, but these can only be used if they have been booked, thus in practice makes them meeting rooms. The reading rooms which is reserved for Master students at IDI is all located in IT-bygget and IT-bygget sydfløy. The one located in IT-bygget sydfløy is called Fiol and is where the researchers seat is located during the master project. The other ones are in IT-bygget. They are Gribb (basement), Sule (2nd floor) and Korp (3rd floor). Although the researchers seat is in Fiol, the researcher is also able to access the other reading rooms as well with the access card. Although the rooms reserved for master student is classified as a reading room, they function more as a computer lab as students is able to set up computers which is borrowed to them from the department. Especially since there are outlets at each desk for both power cables and Ethernet cables.

3.4.1.2 Room Selection

Lecture halls

Although the researcher have access to all lecture halls at Gløshaugen, doing measurements in all would take a long time. Therefore, the researcher has decided to try to measure a few lecture halls from each major building at campus. Choosing which lecture halls from each building to do measurements in have been done by looking at the schedule for each room and noting down the time there are scheduled lectures. If a lecture is scheduled the researcher is more inclined to do measurement for the room, but doing measurements on an empty room is also possible. When the researcher have a list of rooms and the time when there are lectures in them, the researcher tries to set up a personal schedule from Monday to Friday where the aim is to attend many various lectures in different lecture halls while not visiting the same room more than once.

The usual time when lectures are schedules is during the day from 10:00 to 16:00. Almost all lectures span over the default two periods. Such as 10:15 to 12:00, 12:15 to 14:00 and 14:15 to 16:00. For every hour there is a default of 15 minute break time, starting from the top of the hour to 15 minutes past. Given that most lectures are hold from 10:00 to 16:00, it is possible to do up to three measurements a day given that lectures span two periods and one measurement also span the same time. In some measurements the researcher stays in the room after a lecture is finished so that data can also be gathered for when the room is empty (except the researcher). The purpose of this is to see a difference in the result when a room is empty or filled with people.

Table 3.1 shows the lecture halls the researcher has chosen to do measurements in. The date and time of the measurement is also provided. Most of the measurements takes two hours, the same time as a lecture if the break time in front of the lecture is included. Some measures last three hours, for instance in EL6 and R8. Staying an extra hour or arriving one hour earlier before scheduled lectures makes it possible to do measurements when the room is in use and when it is empty.

Reading rooms and computer labs

The reading rooms and computer labs the researcher has chosen to conduct measurements in are those the researcher has access to. These rooms are the two on the first floor of Realfagsbygget, B1-101 and B1-105 (the room B1-165 is also available for the researcher, but as the room is not listed in the Analytics API it is not possible to get device usage data of the room), and the rooms reserved for master students at IDI, Fiol, Gribb, Sule and Korp. As the primary work location for the researcher is in Fiol, the researcher has decided to do more measurements in Fiol as it is convenient for the researcher to do measurements while doing other work.

Table 3.2 shows the reading rooms and computer labs measurements have been performed in.

Room	**Building**	**Date and time**
R7	Realfagsbygget	2016.04.13 10:00 to 12:00
H3	Hovedbygget	2016.04.14 10:00 to 16:00
S3	Sentralbygget	2016.04.15 10:00 to 12:00
S5	Sentralbygget	2016.04.18 10:00 to 12:00
EL5	EL-bygget	2016.04.18 12:00 to 14:00
S1	Sentralbygget	2016.04.19 12:00 to 14:00
EL6	EL-bygget	2016.04.20 10:00 to 13:00
R8	Realfagsbygget	2016.04.20 13:00 to 16:00
R9	Realfagsbygget	2016.04.21 10:00 to 12:00
F2	Gamle Fysikk	2016.04.22 10:00 to 12:00
F6	Gamle Fysikk	2016.04.22 12:00 to 14:00

Table 3.1: Lecture halls the researcher has chosen for conducting measurements in. The table is sorted by the column *Date and time*.

Room	**Building**	**Date and time**	**Room type**
Fiol	IT-bygget sydfløy	2016.04.01 14:15 to 16:15	Computer lab
Fiol	IT-bygget sydfløy	2016.04.04 13:30 to 15:30	Computer lab
Fiol	IT-bygget sydfløy	2016.04.11 13:00 to 17:00	Computer lab
Gribb	IT-bygget	2016.04.25 10:00 to 12:00	Computer lab
Sule	IT-bygget	2016.04.25 12:00 to 14:00	Computer lab
Korp	IT-bygget	2016.04.25 14:00 to 16:00	Computer lab
B1-101	Realfagsbygget	2016.04.26 10:00 to 12:00	Reading room
B1-105	Realfagsbygget	2016.04.26 12:00 to 14:00	Reading room

Table 3.2: Reading rooms and computer labs where measurements have been conducted. The table is sorted by the column *Date and time*.

3.4.1.3 Measurement Strategy

Collecting data about how many people use a room is done by writing down how many people are present in a room at a set time interval. Prior to conducting a measurement for a room the researcher prepares a sheet of paper with a table containing all the points in time for when the researcher have to note down the number of people in the room. Tools, such as alarms and timers, can be used to remind the researcher about when the points in time for writing down the number of people occurs. This ensures that no data point is skipped during a measurement due to distractions.

Counting how many people are in a room can be done in two ways. The first way is that the researcher does a full count of the room at every time interval, and then noting the number down on the sheet of paper. This method can be used when the researcher has a clear view of the whole room and its occupants. This method becomes more difficult the more people are present in a room. For instance, many of the lecture halls at NTNU can hold around 200 peoples. Counting that many people does take time, and during this time people can move around within the room, enter the room or exit the room. This makes it challenging for the researcher to do an accurate count of the room when there are many people in the room.

The second way of counting how many people use a room is that the researcher does an initial count of the room upon arrival, then the researcher monitors the entrances and exits of the room while keeping track of how many people enters or leaves. This requires that the researcher has a clear line-of-sight to the doors of the room. Finding the total count of people in the room using this approach is as simple as adding to the initial count when people enters and subtracting when people leaves. Thus, instead of doing a recount of the room at every time interval, many smaller counts are performed continuously throughout the measurement when someone enters or exits the room. This method provides less uncertainty as counting small groups of people is both easier and faster than counting a large group. In cases where the researcher loses track of the total count when using this method, the counting method can be restarted by doing a new recount of the entire room, and then counting how many enters and exits. Situations where the researcher can lose track of the count is when many people leaves the room at the same time, such as when students leaves the lecture hall during break time for lectures, or when a room has two or more doors with people entering or leaving at the same time. It is also possible to measure a room from the outside if it only has a single entrance/exit, but if the researcher loses count, one must enter the room to do a full room count.

As the Analytics System is able to provide device count data at a granularity of once every minute, there are many options when choosing which interval to use when conducting measurements. For a two hour measurement, using a time interval of every 15 minutes would result in 9 data points, every 10 minutes result in 13 data points, and every 5 minutes result in 25 data points. Using an interval smaller than every 5 minutes is possible, but the room count does usually not change much that using a 5 minute interval gives a worse representation of how many people were in the room. Situations where the room count changes quickly is when a lecture is about to start, during break time of a lecture, and when a lecture is over. A lot of changes can occur when a lecture ends and people are leaving at the same time as new students enters for the lecture in the room.

The break time during lectures lasts 15 minutes. Using a time interval of every 15 minutes would give a bad representation of how many people was in the room during the break. Visualized on a graph this could result to a single data point being much lower when the break occurred. It could also happen that no change is detected if the time the researcher writes down the room count is right before the break starts, and then the next time to write down the room count students have reentered the room. Using a time interval of every 10 minutes could also result in either one or two data point representing the state of the room during the break. Choosing a 5 minute time interval would give better resolution for getting data points showing the room count during a break.

In reading rooms and computer labs there is no default break time for students. People come and leave when they want. Therefore it is not necessary to have a time interval as low as every 5 minute when doing the measurements as there are no fixed fixed times when people leaves or enters. If the researcher is focused on only doing the measurement and not other work, there are no downsides to using a time interval of every 5 minutes. In situations where the researcher is doing work unrelated to the measuring it is possible to choose a larger interval. A time interval of every 15 minutes had been chosen in the measurements performed in Fiol.

Another strategy that can be used when conducting measurements is to measure two rooms at the same time. This is possible if the researcher is located somewhere with line-of sight to both of the- entrances of the rooms that would be measured. It is important that the rooms does not also have entrances in other locations where the researcher is not able to see. When the researcher is able to see both entrances it is easy to keep track of how many people enters or leaves the rooms, but it is necessary that the researcher does a full count of the rooms when starting the measurement. If the researcher loses the count in any of the rooms it would be hard to do a recount as that would mean entering one of the rooms and spend some time to do the counting, which makes it possible that people could enter or leave the other entrance. Therefore, doing a measurements where two rooms is measured at once is difficult as it is important that the researcher manages to stay focused and not lose count in one of the rooms.

3.4.1.4 Measurements Accuracy

The accuracy of the number of people that the researcher is able to count in a room depends primarily on how many people are in the room. In general, the more people there are in a room the harder it is to do an accuracy full count of the room as it takes time to count everyone and people could be moving around, entering or exiting the room. Therefore the accuracy of the measurement is better if the researcher manages to count by keeping track of how many people enters or leaves the room through a rooms single exit. If there are more than one doors where students can use, it is often difficult to keep track of how many leaves or enters the room, especially during break times for lectures.

Reading rooms and computer labs have a small capacity and therefore are easy to keep track of all the persons in it. If the researcher manages to lose track of the current count in the room it is easy to take a quick walk through the room and get an accurate count of the number of people. Some lecture halls are small which makes it easy to do full counts. Examples of such lecture halls is F2 and F6, where the former has a capacity of 72 and the latter 88.

Another factor which makes it hard to do full counts in lecture halls is the shape of the room. If the seats are placed in a plane instead of a curve it is hard for the researcher to get a clear line of sight throughout the room as sight is obstructed by people, making it hard for the researcher to see what is behind people. The more people there are the more of the room will be obstructed. If the researcher is able to monitor the single entrance to a room in order to keep track of the number of people, the shape of the room will not affect the measured count much. In the cases where a lecture hall is already filled up with students when the researcher enters the room prior to starting measurements, it could be hard to get an accurate full room count when the seats are placed in a plane, but the researcher can often stand up while doing these counts allowing for a better line-of-sight

Given the factors discussed in the above paragraphs, the accuracy of the number of people on a room is not always 100% correct. The more people are present in a room the larger the margin of error becomes as it is more difficult to keep track of everyone in the room at all times. For instance: Given a room with a large capacity has about 150 persons currently in the room, it is reasonable to believe that the researcher will be able to get a count which is close to the actual number of people, 150, but due to errors in counting the researcher could have counted somewhere between 145 - to 155 people. If the researcher has an error of ±5 people in a room with 150 people, the actual number would likely not be the number the researcher is able to count. If it proves that the accuracy of the Analytics System is very good at assessing how many devices there are in a room based on the number of people present, it would be important to measure exactly how large the margin of error is when the researcher count the number of people in the room. But given that there already is an uncertainty in the positioning data of 5-10 meters [Biczok et al., 2014], it is unlikely that some uncertainty in the counting of the number of people in a room would affect the research negatively.

3.4.2 Calculated Device Count

The Analytics System (see section 2.2.3) uses the data collected by the Cisco MSE (see section 2.2.2) in order to calculate how many devices are located in each room on campus Gløshaugen. Obtaining this data requires the researcher to have access to the Analytics API with a username and password combination in order to obtain a necessary session key. Without it the API will not return any data when using the URL's which is described below. When entering `analytics.mazemap.com` in a browser, which is the application developed by Mersland [2013], a login screen is shown if there is currently no session set up to the Analytics API. After successfully logging in the session will be stored as a cookie, and any further URL-requests to the API sent through the same browser will receive valid results.

3.4.2.1 List of Rooms

In order to get calculated device count for a room from the Analytics API it is necessary to know the ID of the room. This ID can be found using an URL which can return a list of rooms. The URL for returning a list of all lecture halls in the system is:

- `https://analytics.mazemap.com/api/campus/1/regionlists/1/`

Returning lists of other room types is also possible. Changing the digit in the last bit `/1/` to either 2, 3 or 4 will return other room types. An overview of which room type is returned when changing the last digit is as follows:

- `.../1/` - Lecture halls
- `.../2/` - Reading rooms
- `.../3/` - Meeting rooms
- `.../4/` - Computer labs

The result when sending one of the URLs above is a JSON-object (JavaScript Object Notation) containing a list of rooms of the selected type. An excerpt of a room object and what information it contains is shown in listing 3.1

```
{
    "campusId": 1,
    "capacity": null,
    "floorIdent": "318",
    "geometries": [
        ... // Omitted due to number of lines.
    ],
    "id": 14820,
    "identifierId": 38312,
    "itemIdent": "108",
    "name": "Fiol",
    "predictionGroupId": 1
}
```

Listing 3.1: Example of the structure of a room object when requesting a list of rooms from the Analytics API. The room Fiol, which is located in IT-bygget sydfløy is used in the example.

There are a many different attributes for each room object. For instance, *campusId* refers to the ID of the campus. Gløshaugen has a value of *1*. The *capacity* is most often set to *null* for most rooms as Wireless Trondheim does not have that information for all rooms. Some lecture halls have the capacity set as it is possible to find out how many seats there are in such a room. The attributes *floorIdent* refers to the id of the floor the room is on. The *geometries* attribute contains a list of coordinates which describes the boundaries of the room. The coordinates uses the longitude and latitude format, which allows the system to know the exact shape of the room as it is in the real world. *IdentifierId* is used internally by the Analytics System to keep track of unique IDs of different object.

The two most important attributes is the *id* and *name*, whereas the latter is the name of the room. Some rooms does not have a name such as Fiol but does instead consist of characters and letters, such as B1-105 which is a reading room located in Realfagsbygget. The researcher can get the ID of a room by searching through the JSON-object for the name of the desired room. The system does contain most rooms, but some rooms are missing, such as F1 which is one of the largest lecture hall on campus Gløshaugen.

3.4.3 Usage Data for Room

In order to request device usage data from a room it is necessary to know the ID of the room. Obtaining the ID was explainied in the previous section. With the ID at hand, such as the one in the previous listing example 3.1 for Fiol, the base URL when requesting data for the room is as follows:

- `https://analytics.mazemap.com/api/campus/1/regions/14820/usagegraph/`

Note that the *id*, 14820, from listing 3.1 appears in the URL. Sending the above URL will result in a *Bad Request*-response. This can be fixed by adding parameters at the end of the URL. Most parameters are optional, but one is required. The required one is the *from* parameter. This tells the system the start of the time interval for data to be returned. Other parameters, which are optional, are *to* and *points*. Following is a summary of the parameters and how they affect the result:

- **from** (**Required**) - Specifies the start point of the time interval for usage data to return.
- **to** (**Optional**) - Specifies the end point of the time interval for usage data to return. Must be at a later point in time that the *from* parameter. If *to* is not specified the system assumes *to* to be the current time of the system.
- **points** (**Optional**) - Specifies how many data points will be returned from the time interval. There exists data points for every minute in the system. If the time interval between *from* and *to* is 120 minutes and *points* is set to 60, the system will only return 60 different usages from the total of 120 minutes (since the system stores usage for every minute). The system attempts to spread the returned points evenly across the specified time interval.

The *to* and *from* parameters is of the form *Epoch* time, or *Unix* time. More specifically the Analytics System uses the Epoch time in milliseconds to store dates. An example of an URL with parameters that obtains all data points from the measurements done in Fiol on 2016.04.01 (see section 4.2.1) looks like this:

- `https://analytics.mazemap.com/api/campus/1/regions/20503/usagegraph/?from=1459512900000&to=1459520100000&points=300`

The measurement in Fiol on 2016.04.01 was from 14:15 and lasted to 16:15, a two hour session. In order to find out the equivalent epoch times for the *from* and *to* parameters, a tool such as EpochConverter[1] can be used. The tool works by typing in the date: year, month, day, hour, minute and second. It is also important to have the correct time zone selected when converting to epoch timestamp. Selecting *Local time* from the options of timezones ensures that it is the same timezone as Analytics System uses.

As the researcher requests usage data for a period of two hours (14:15 to 16:15) it is necessary to have a point value equal to or higher than the number of minutes in the time interval. Two hours including the interval bounds equals to 121 minutes (2 times 60 minutes plus 1 minute). Setting the points value to 300, which is sufficient for up to 5 hours, assures that the system will return all usage data for the specified time interval.

The result when sending the above URL to the system is a JSON-object which is a list-object consisting of *usages*. Each usage looks like the one in listing 3.2. There are only three attributes, or fields, in this object. *Id* is a unique ID given to each usage object. *TickTime* is the epoch time for the usage. When converting

---

[1] http://www.epochconverter.com

the value in the listing to a human readable format, the result is 4/1/2016, 2:17:00 PM GMT+2:00 DST (American Month/Day/Year-format). The last field is how many devices the Analytics System has calculated to be in the room at the given timestamp.

```json
{
    "id": 1149181,
    "tickTime": 1459513020000,
    "usage": 3.00267
}
```

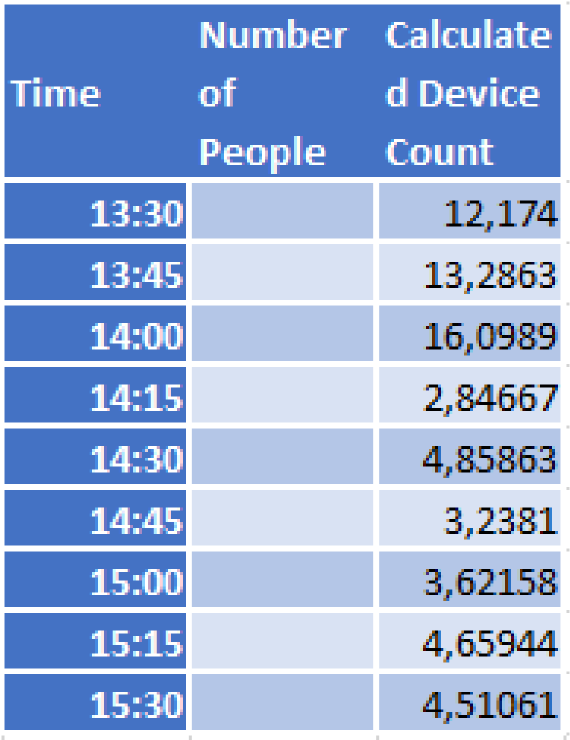
Listing 3.2: Example of a usage data-object (JSON format) which is one of many usage objects returned when requesting usage data for a specific room from the Analytics System.

When the researcher gets a list of all the usages from the same time interval a measurement has been performed, it is not often needed to use all the usage data which is returned as the researcher is often interested in the usages with a *timeTick* which matches the time interval used in the measurement. For instance every 5, 10 or 15 minute intervals where the researcher notes down the number of people in the room. The researcher did create a small Java-program which takes the whole list of usages as input, and outputs a filtered list where only the usages that matches a desired time interval is outputted. The output is also formatted such that the Epoch time is converted to human readable form. It is possible to take the output from the program and copy and paste it into Excel such that three columns is created automatically. The first column will contain the timestamps for each usage. The second column is empty but is meant for the researcher to manually enter the data gathered during a measurement (number of people). Lastly, the third column have all the usage data already created, where the usage matches the timestamp in the first column. Using the Java-program saves the researcher a lot of time as it is not needed to manually copy and paste each usage from each usage object in the list of usages. An example of how the data will look in Excel after basic formatting is applied to distinguish the column headers and the first column is shown in figure 3.3.

This report is written in LaTeX. In order to convert a table from Excel to a format which can be entered into the Latex-document without too much copy and paste, the researcher had used a tool found at `http://www.tablesgenerator.com/`. By simply copying the table from excel and pasting it into the tool, it is possible to click on *generate* in order to convert the whole table into a Latex form which can easily be pasted into this report. When conducting many measurements and getting a lot of data it would be time consuming to create tables and inserting the data manually.

3.5 Data Analysis

It is possible to compare the two types of data, number of people (from measurements) and calculated device count (from Analytics System), by plotting both data



Time	Number of People	Calculated Device Count
13:30		12,174
13:45		13,2863
14:00		16,0989
14:15		2,84667
14:30		4,85863
14:45		3,2381
15:00		3,62158
15:15		4,65944
15:30		4,51061

Figure 3.3: The output from the Java-program is ready to be copy and pasted into excel in order to create a structured table. In the figure the column headers and the first column has been formatted after the researcher had pasted the table into Excel. What is left for the researcher is to enter the number of people data gathered during a measurement.

in a graph and look at the data plots to see if there is a correlation between the two data sets. A more statistically approach is to use mathematics in order to determine how much correlation there is between two data sets.

3.5.1 Pearson Product-Moment Correlation Coefficient

The Pearson Product-Moment Correlation Coefficient (PPMCC) is a measure of the linear correlation between two variables X and Y. The resulting correlation is often denoted as r. It is also possible to calculate the correlation on two data sets. Assuming we have one data set $\{x_1, ..., x_n\}$ containing n values and another data set $\{y_1, ..., y_n\}$ also containing n values, then the formula [Wikipedia] for r is as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

where:

- n, x_i, y_i are defined as above
- $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; and analogously for $\overline{y}$

Using the above formula manually to compare many and large data sets would take time. Instead it is possible to use an inbuilt function in Excel to do the work automatically. The *correl* function can do it automatically if the two data sets are provided as parameters to the function. For instance: If the first data set is in column B and the second data set in column C, where the first data entries are in row 2 and the last in row 10, the following function will calculate the PPMCC:

- **=CORREL(B2:B10;C2:C10)**

Note that colon is used between B2 and B10 to denote a range for the data set to be all cells between and including B2 and B10. Semicolon is used to differentiate between parameters.

Chapter 4

Research Result

This chapter presents the results from the measurements conducted throughout the project. For each measurement information about the room is provided. This is called the metadata, and contains information of room name, building, room capacity, date and time of the measurements as well as information about what the schedule for the room was during the measurement. After the metadata is a figure which shows where the room is located on Gløshuagen. Following the figure is the data from the measurement presented in a table, as well as additional details regarding the data collection. Both the *Number of People* and *Calculated Device Count* data is presented in the table. Following the table is a graph which visualizes the data from the table. After the graph follows the calculation of correlation between number of people and calculated device count. The Pearson Product-Moment Correlation Coefficient, PPMCC, is used on the data from the table to calculate the correlation. Lastly, notes from when the measurement was conducted is shown. These are notes made by the researcher during the measurement.

Lecture halls will be presented first, then follows the computer labs and reading rooms. Lastly there is a description of other planned measurements and reasons for why they were not conducted. It includes information regarding the Analytics System from Håkon Lorentzen, the man which is currently responsible for the system at Wireless Trondheim.

4.1 Lecture Halls

The measurements for lecture halls is listed by date and time. This means they are presented in the order the measurement was done.

4.1.1 Measurement on 2016.04.13 in R7

4.1.1.1 Metadata

Information regarding the measurement is in table 4.1 below.

Room	R7
Building	Realfagsbygget
Capacity	342
Date	2016.04.13
Time	10:00 to 12:00
Schedule	Lecture (KJ1020 – Organisk Kjemi)

Table 4.1: Metadata for measurement in R7 on 2016.04.13.

4.1.1.2 Map Location

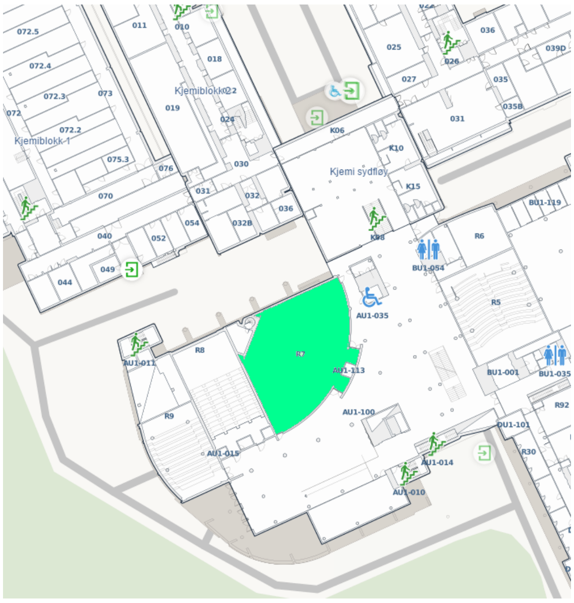


Figure 4.1: Map showing where R7 is located. The room is highlighted in green.

4.1.1.3 Data

Data collected throughout the measurement is in table 4.2. As this was the researchers first measurement in a lecture hall, the chosen interval to note down room count was every 15 minutes, and strategy for counting the room was to do a whole room count every interval. As there was many people in the room it took a long time to count everyone. About halfway through the measuring, at 10:54, the researcher decided to change strategy by counting only how many entered and leaved the room. This was not possible through the long break from 11:00 to 11:21 as the room has two doors which made it difficult to keep track of how many people entered or left the room during the break. After the break the researcher was able to note down the count of people at an interval of every 5 minute as counting by monitoring how many people entered of left the room is quicker and easier than doing a full count at every time interval. Also note that the first count was when the researcher was able to enter the room at 10:08.

Time	Number of People	Calculated Device Count
10:08	82	21.2131
10:15	138	20.0855
10:30	153	28.5465
10:45	156	2.90981
11:00	77	14.9624
11:18	155	11.3785
11:20	157	8.96914
11:25	157	5.48298
11:30	157	20.4803
11:35	157	9.59415
11:40	157	20.2857
11:45	157	4.05296
11:50	157	11.9794
11:55	157	10.0126
12:00	157	13.734

Table 4.2: Data collected for measurement in R7 on 2016.04.13.

4.1.1.4 Graph

The graph from the data in table 4.2 is shown in figure 4.2.

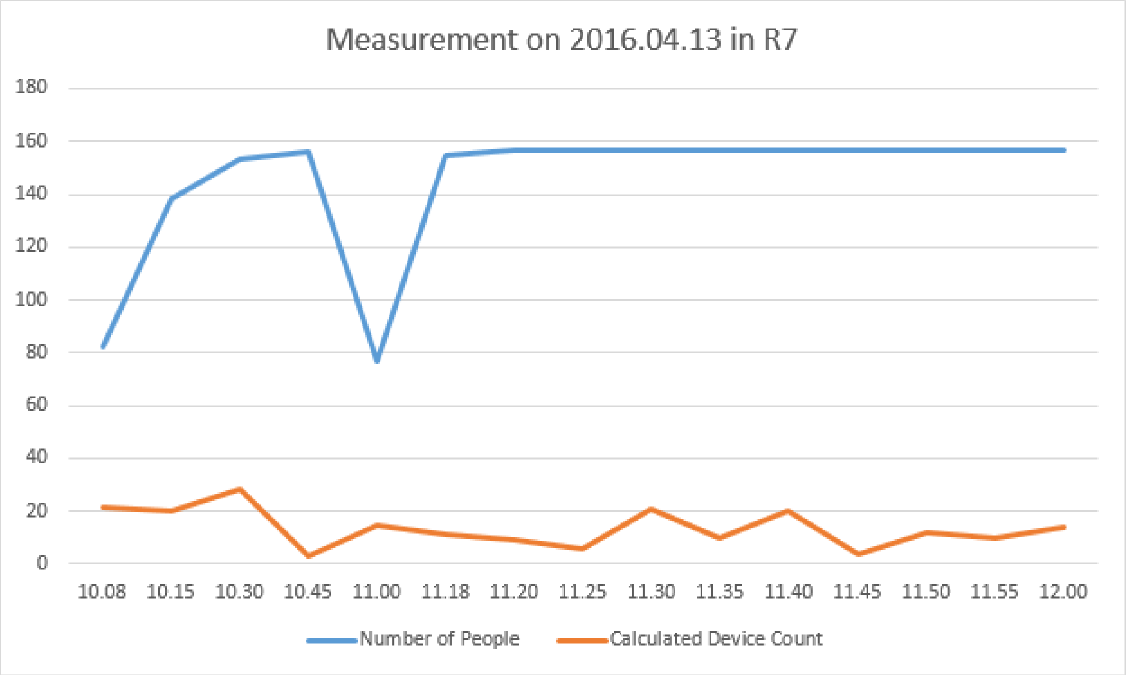


Figure 4.2: Graph visualizing the data from measurement on 2016.04.13 in R7.

4.1.1.5 Correlation

The PPMCC between the two data sets *Number of People* and *Calculated Device Count* from table 4.2 is: -0.30827.

4.1.1.6 Notes

These are notes taken by the researcher during the measurement.

- Can't see any laptops in use. Students take notes by hand and use calculators.
- It is difficult to do measurements at break time as there are several doors for students to enter or leave the room.

4.1.2 Measurement on 2016.04.14 in H3

4.1.2.1 Metadata

Information regarding the measurement is in table 4.3 below.

Room	H3
Building	Hovedbygget
Capacity	189
Date	2016.04.14
Time	10:00 to 16:00
Schedule	10:15 to 12:00: Lecture (TKT4102 - Dynamikk)
—”—	12:15 to 14:00: Lecture (TMA4115 - Matematikk 3)
—”—	14:00 to 16:00: No lecture

Table 4.3: Metadata for measurement in H3 on 2016.04.14.

4.1.2.2 Map Location

Figure 4.3: Map showing where H3 is located. The room is highlighted in green.

4.1.2.3 Data

The researcher learned in the previous measurement in R7 (see section 4.1.1) that a time interval at 15 minutes was too long, and decided to try 5 minute interval for this measurement. The measurement strategy was also to monitor the only door for the room in order to track how many people was in the room. As explained in section 2.2.3, the Analytics System does not calculate room count at some times during the day. At the day this measurement was conducted the system did not calculate room count from 12:30 to 13:50. Therefore, two tables for data has been created for the measurement in H3. The first (table 4.4) from 10:00 to 12:25, and the second (table 4.5) from 13:55 to 16:00.

Time	Number of People	Calculated Device Count
10:00	3	0.0
10:05	9	0.0
10:10	20	1.98981
10:15	31	1.54229
10:20	36	5.10613
10:25	39	1.01942
10:30	40	4.72082
10:35	40	0.997647
10:40	41	0.640013
10:45	41	3.57473
10:50	41	0.994904
10:55	41	3.67189
11:00	41	2.53993
11:05	32	1.86011
11:10	30	1.98981
11:15	38	0.0
11:20	38	1.98981
11:25	38	0.0
11:30	38	0.994904
11:35	38	0.0
11:40	38	0.994904
11:45	38	1.98981
11:50	38	0.0
11:55	38	0.797888
12:00	38	0.0
12:05	9	0.364747
12:10	47	1.54229
12:15	83	0.0
12:20	90	1.20463
12:25	92	1.79149

Table 4.4: Data collected for measurement in H3 on 2016.04.14 - Part 1.

Time	**Number of People**	**Calculated Device Count**
13:55	92	2.08967
14:00	92	1.98981
14:05	10	0.994904
14:10	1	0.0
14:15	1	0.0
14:20	1	0.0
14:25	1	0.0
14:30	1	0.0
14:35	1	0.0
14:40	1	0.0
14:45	1	0.0
14:50	1	0.0
14:55	1	0.586791
15:00	1	0.0
15:05	1	0.994904
15:10	1	0.0
15:15	1	0.999441
15:20	1	0.0
15:25	1	0.0
15:30	1	0.0
15:35	1	0.0
15:40	1	1.98981
15:45	1	0.0
15:50	1	0.0
15:55	1	0.0
16:00	1	0.0

Table 4.5: Data collected for measurement in H3 on 2016.04.14 - Part 2.

4.1.2.4 Graph

Since there are two data tables due to the Analytics System not calculating a device count from 12:30 to 13:50, there will be two graphs. The first (figure 4.4) represents the data in the first table (table 4.4), and the second (figure 4.5) represents the data in the second table (table 4.5).

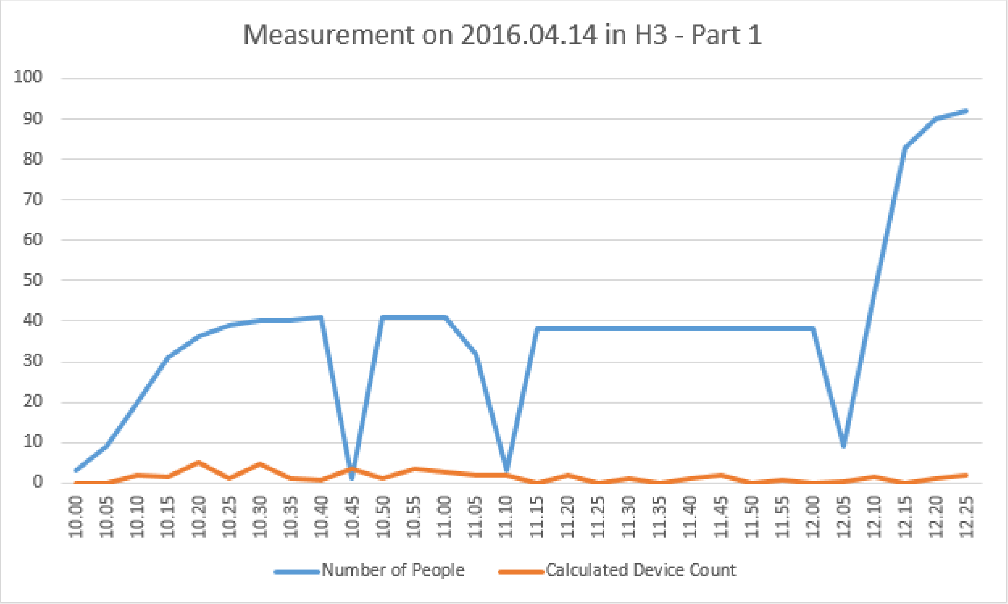


Figure 4.4: Graph visualizing the data from measurement on 2016.04.14 in H3 - Part 1.

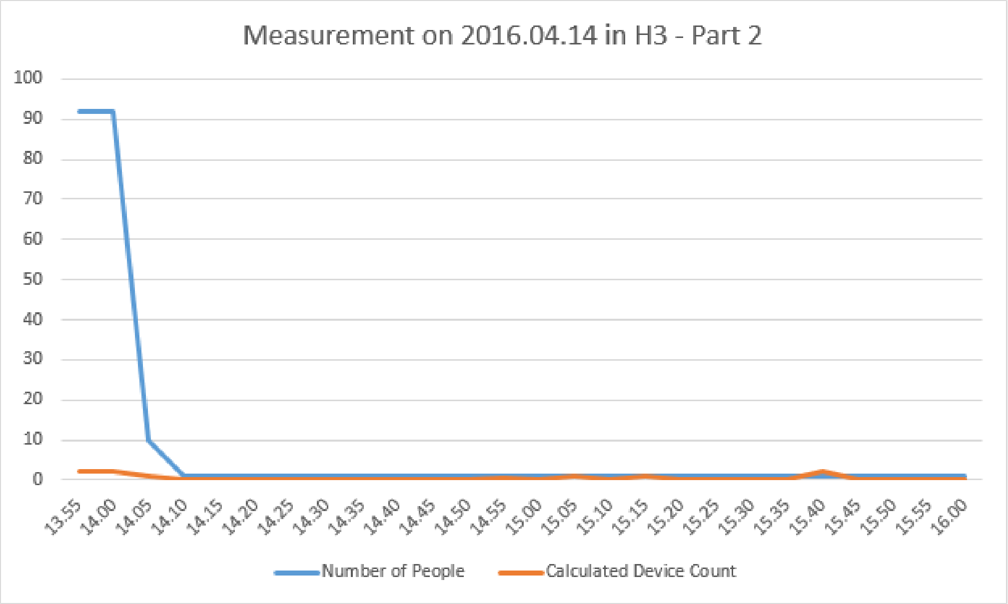


Figure 4.5: Graph visualizing the data from measurement on 2016.04.14 in H3 - Part 2.

4.1.2.5 Correlation

The PPMCC will be calculated for each of the two tables containing data. Table 4.4 gives the value: -0.04742, and table 4.5 gives: 0.722433.

4.1.2.6 Notes

- During the first lecture students took notes by hand from the blackboard. No laptops in use.
- During the second lecture students took notes by hand from the blackboard. No laptops in use.
- No lecture from 14:00 to 16:00. Only the researcher was present.
- A single wall mounted AP in the room.

4.1.3 Measurement on 2016.04.15 in S3

4.1.3.1 Metadata

Information regarding the measurement is in table 4.6 below.

Room	S3
Building	Sentralbygget
Capacity	250
Date	2016.04.15
Time	10:00 to 12:00
Schedule	10:15 to 12:00: Lecture (TEP4100 - Fluidmekanikk)

Table 4.6: Metadata for measurement in S3 on 2016.04.15.

4.1.3.2 Map Location

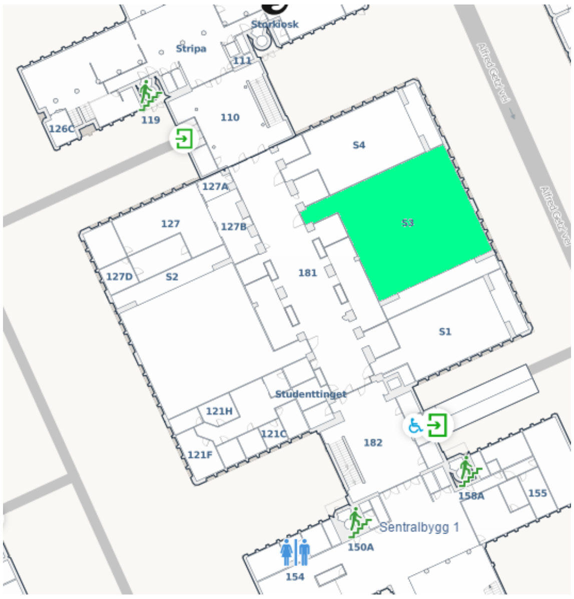


Figure 4.6: Map showing where S3 is located. The room is highlighted in green.

4.1.3.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.7.

Time	Number of People	Calculated Device Count
10:00	5	0.0
10:05	24	0.0
10:10	64	0.0
10:15	107	0.0
10:20	136	0.0
10:25	144	1.9885
10:30	148	2.63324
10:35	150	0.640013
10:40	151	0.321957
10:45	151	1.5242
10:50	151	0.43778
10:55	151	0.0
11:00	151	1.54229
11:05	67	1.67867
11:10	60	0.993599
11:15	80	0.43778
11:20	150	2.46184
11:25	151	0.450264
11:30	151	0.994904
11:35	151	0.640013
11:40	151	1.37564
11:45	151	0.0
11:50	151	0.0
11:55	151	0.994904
12:00	151	0.0
12:05	36	0.0

Table 4.7: Data collected for measurement in S3 on 2016.04.15.

4.1.3.4 Graph

The graph from the data in table 4.7 is shown in figure 4.7.

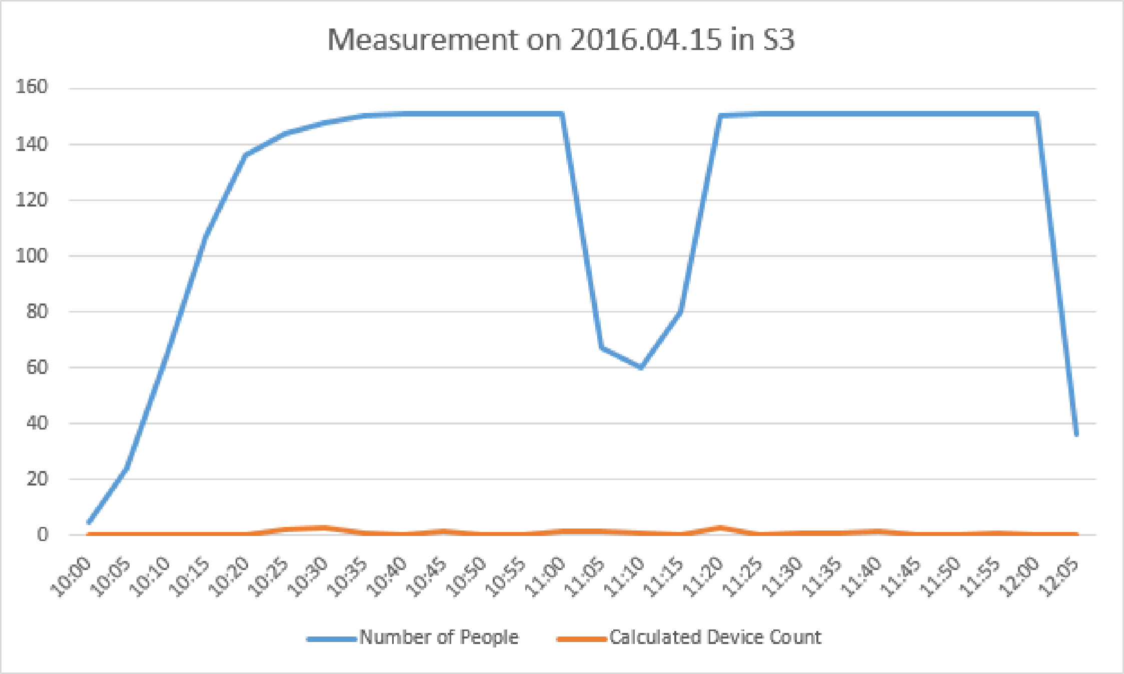


Figure 4.7: Graph visualizing the data from measurement on 2016.04.15 in S3.

4.1.3.5 Correlation

The PPMCC from the data in table 4.7 is 0.302237.

4.1.3.6 Notes

- No laptops in use. Notes were taken by hand from blackboard.

4.1.4 Measurement on 2016.04.18 in S5

4.1.4.1 Metadata

Information regarding the measurement is in table 4.8 below.

Room	S5
Building	Sentralbygget
Capacity	222
Date	2016.04.18
Time	10:00 to 12:00
Schedule	10:15 to 12:00: Lecture (TBA4100 – Geoteknikk og Geologi)

Table 4.8: Metadata for measurement in S5 on 2016.04.18.

4.1.4.2 Map Location

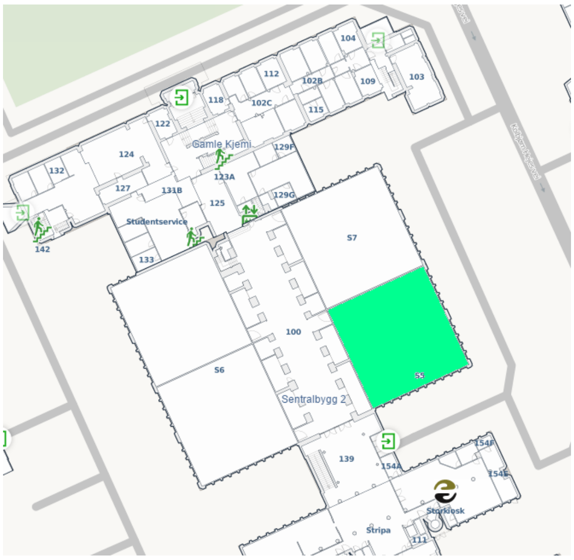


Figure 4.8: Map showing where S5 is located. The room is highlighted in green.

4.1.4.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.9.

Time	**Number of People**	**Calculated Device Count**
10:05	28	6.76117
10:10	71	5.90086
10:15	105	9.78789
10:20	119	5.3491
10:25	124	10.2574
10:30	126	1.37246
10:35	127	11.188
10:40	127	5.03813
10:45	129	11.3197
10:50	129	6.16936
10:55	129	5.48008
11:00	129	2.64799
11:05	43	4.55865
11:10	48	2.55042
11:15	129	10.7869
11:20	129	4.2853
11:25	129	7.28718
11:30	128	0.987663
11:35	128	2.27796
11:40	128	5.29522
11:45	128	9.8239
11:50	128	4.42988
11:55	128	5.96943
12:00	128	3.08111

Table 4.9: Data collected for measurement in S5 on 2016.04.18.

4.1.4.4 Graph

The graph from the data in table 4.9 is shown in figure 4.9.

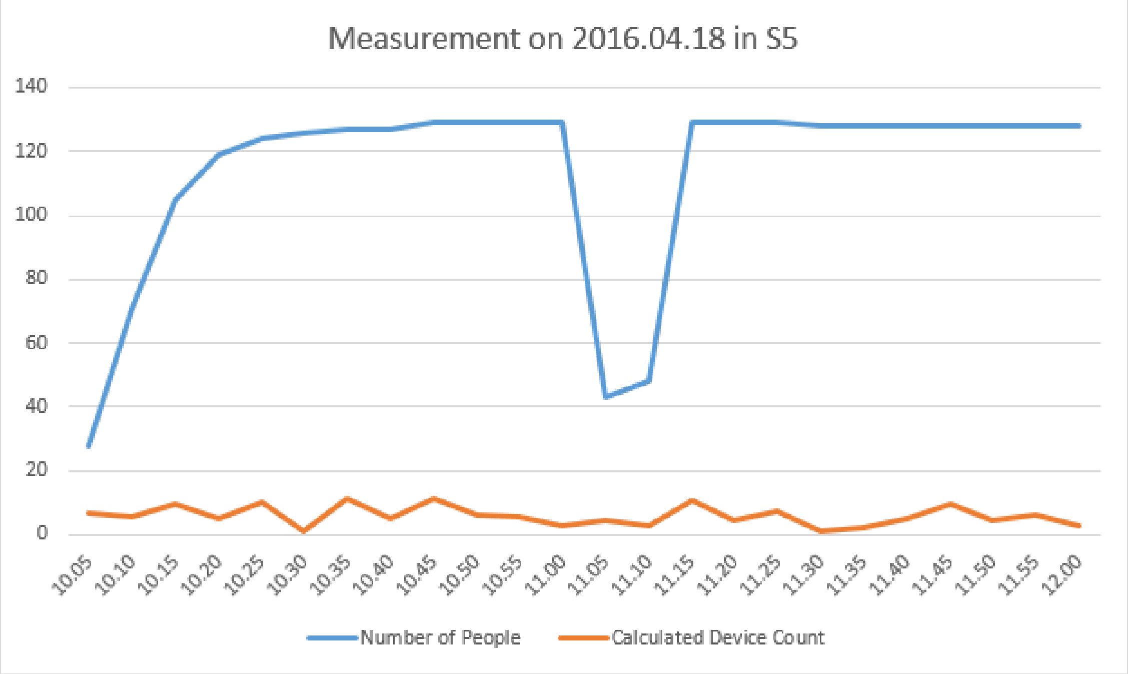


Figure 4.9: Graph visualizing the data from measurement on 2016.04.18 in S5.

4.1.4.5 Correlation

The PPMCC from the data in table 4.9 is 0.097469.

4.1.4.6 Notes

- No laptops in use. Notes were taken by hand from blackboard.

4.1.5 Measurement on 2016.04.18 in EL5

4.1.5.1 Metadata

Information regarding the measurement is in table 4.10 below.

Room	EL5
Building	Gamle Elektro
Capacity	300
Date	2016.04.18
Time	12:00 to 14:00
Schedule	12:15 to 14:00: Lecture (TMM4100 - Materialteknikk)

Table 4.10: Metadata for measurement in EL5 on 2016.04.18.

4.1.5.2 Map Location

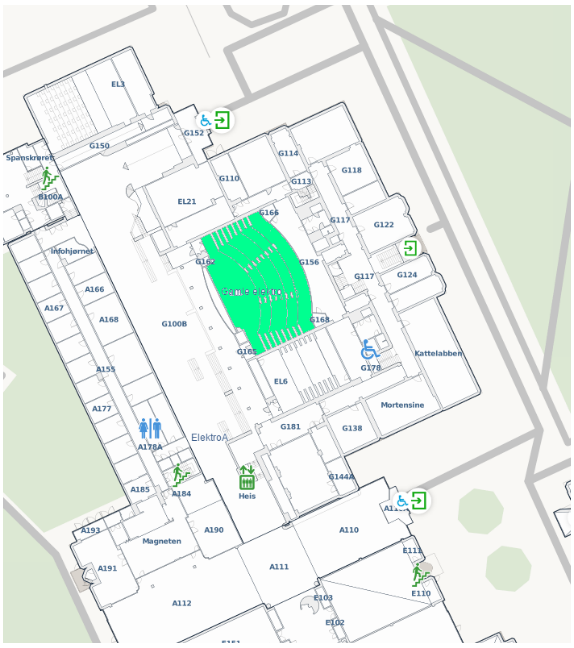


Figure 4.10: Map showing where EL5 is located. The room is highlighted in green.

4.1.5.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.11.

Time	Number of People	Calculated Device Count
12:05	52	1.48367
12:10	70	0.0
12:15	131	0.264834
12:20	131	1.15959
12:25	157	0.0
12:30	158	0.0
12:35	157	4.55598
12:40	157	2.76523
12:45	157	7.12504
12:50	156	4.43723
12:55	156	0.0
13:00	73	0.0
13:05	54	0.0
13:10	57	0.994904
13:15	102	0.0
13:20	146	1.98981
13:25	147	4.75487
13:30	147	3.29572
13:35	147	1.2495
13:40	148	8.40064
13:45	148	9.83234
13:50	148	3.36227
13:55	148	10.4504
14:00	15	4.14661

Table 4.11: Data collected for measurement in EL5 on 2016.04.18.

4.1.5.4 Graph

The graph from the data in table 4.11 is shown in figure 4.11.

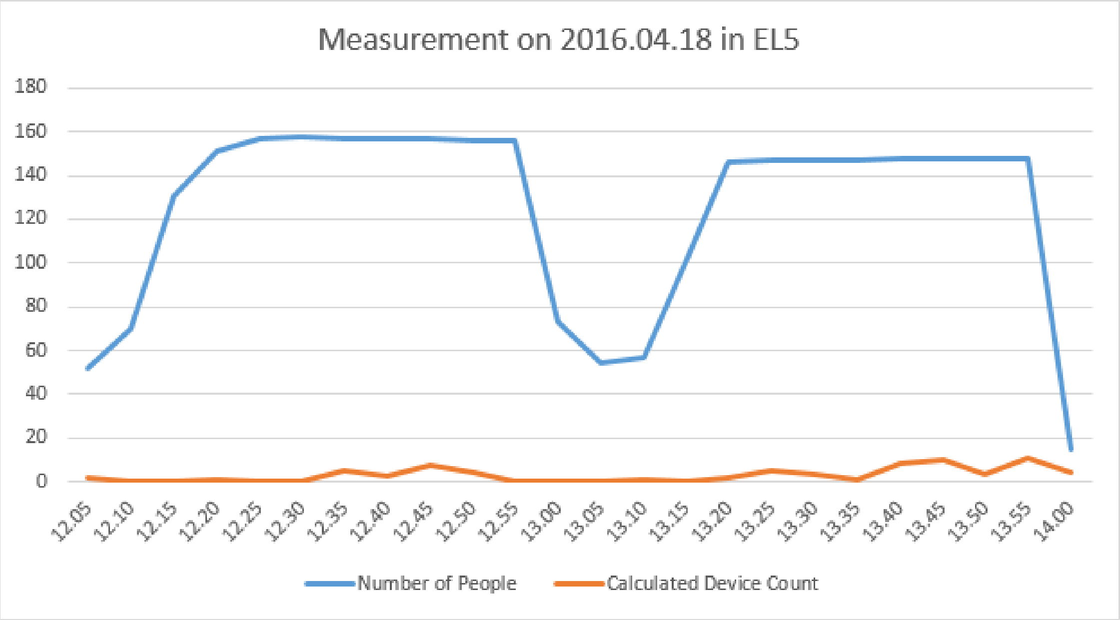


Figure 4.11: Graph visualizing the data from measurement on 2016.04.18 in EL5.

4.1.5.5 Correlation

The PPMCC from the data in table 4.11 is 0.313457.

4.1.5.6 Notes

- No laptops in use. Notes were taken by hand from blackboard.

4.1.6 Measurement on 2016.04.19 in S1

4.1.6.1 Metadata

Information regarding the measurement is in table 4.12 below.

Room	S1
Building	Sentralbygget
Capacity	114
Date	2016.04.19
Time	12:00 to 14:00
Schedule	12:15 to 14:00: Lecture scheduled, but room was empty

Table 4.12: Metadata for measurement in S1 on 2016.04.19.

4.1.6.2 Map Location

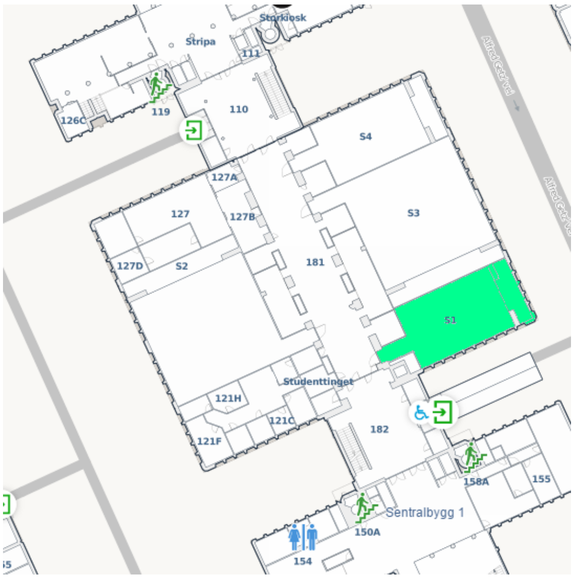


Figure 4.12: Map showing where S1 is located. The room is highlighted in green.

4.1.6.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.13.

Time	**Number of People**	**Calculated Device Count**
12:00	0	1.64429
12:05	2	0.0
12:10	1	0.0
12:15	1	0.116563
12:20	1	0.900529
12:25	1	0.157979
12:30	1	0.20043
12:35	1	0.0
12:40	1	0.450264
12:45	1	0.450264
12:50	1	0.334986
12:55	1	0.0
13:00	1	0.0
13:05	1	0.0
13:10	1	0.0
13:15	1	0.0
13:20	1	0.994904
13:25	1	0.0
13:30	1	1.18062
13:35	1	1.44517
13:40	1	0.157979
13:45	1	0.294805
13:50	1	0.116563
13:55	1	2.0367
14:00	2	0.483759

Table 4.13: Data collected for measurement in S1 on 2016.04.19.

4.1.6.4 Graph

The graph from the data in table 4.13 is shown in figure 4.13.

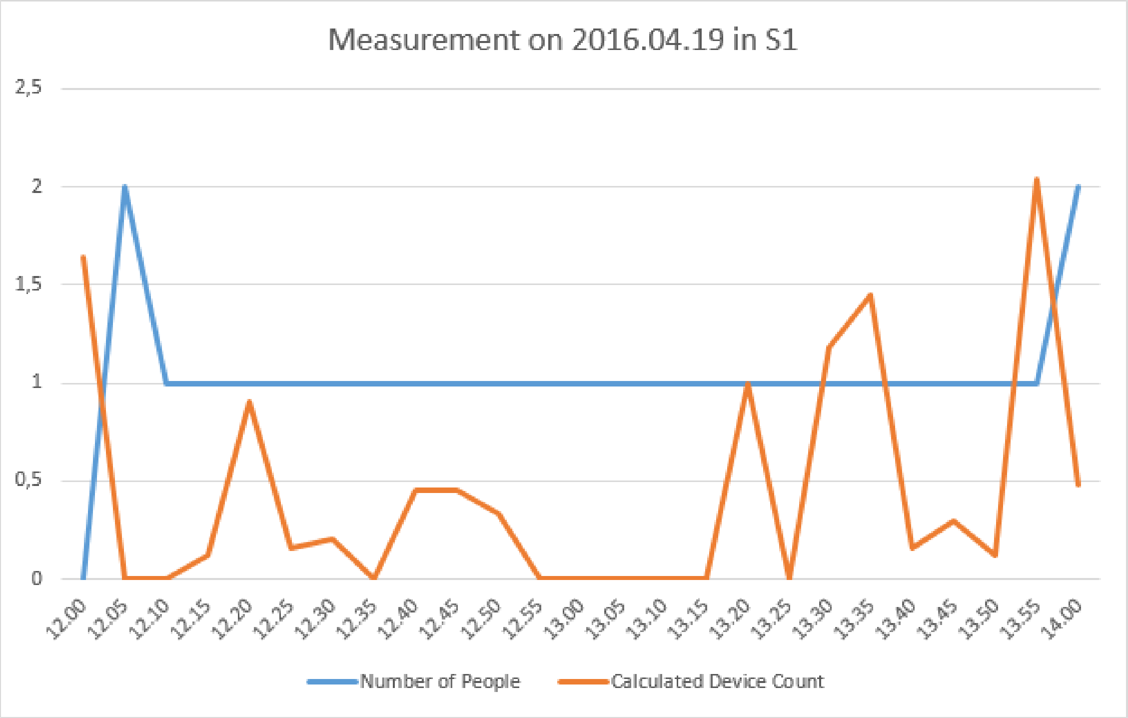


Figure 4.13: Graph visualizing the data from measurement on 2016.04.19 in S1.

4.1.6.5 Correlation

The PPMCC from the data in table 4.13 is −0.323.

4.1.6.6 Notes

- There was a lecture scheduled, but the room was not used. The researcher was present in the room all the time except for the first measure time at 12:00 where the researcher was located right outside the door to the room.
- 1 AP visible in the room.

4.1.7 Measurement on 2016.04.20 in EL6

4.1.7.1 Metadata

Information regarding the measurement is in table 4.14 below.

Room	EL6
Building	Gamle Elektro
Capacity	130
Date	2016.04.20
Time	10:00 to 13:00
Schedule	10:15 to 12:00: Lecture (MA1202 - Lineær Algebra med Anvendelser)
—”—	12:15 to 13:00: No lecture.

Table 4.14: Metadata for measurement in EL6 on 2016.04.20.

4.1.7.2 Map Location

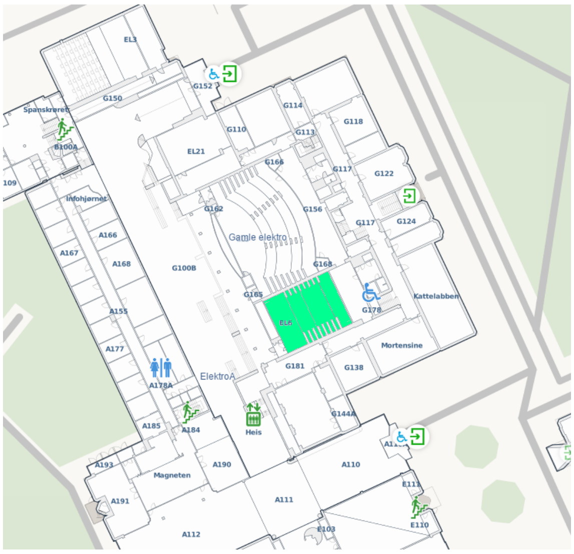


Figure 4.14: Map showing where EL6 is located. The room is highlighted in green.

4.1.7.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.15.

Time	Number of People	Calculated Device Count
10:05	8	0.0
10:10	18	0.450264
10:15	36	0.0
10:20	42	0.0
10:25	44	0.0
10:30	44	0.0
10:35	44	0.0
10:40	44	0.0
10:45	44	0.0
10:50	44	0.166993
10:55	44	0.0
11:00	40	0.412237
11:05	22	0.291989
11:10	25	0.0
11:15	43	0.0
11:20	42	0.0
11:25	43	0.0
11:30	43	0.237028
11:35	43	0.450264
11:40	43	0.0
11:45	43	0.0
11:50	43	0.0
11:55	43	0.0
12:00	43	0.0
12:05	1	0.0
12:10	1	0.0
12:15	1	0.0
12:20	1	0.0
12:25	1	0.310567
12:30	1	0.0
12:35	1	0.0
12:40	1	0.0
12:45	1	0.356221
12:50	1	0.450264

Table 4.15: Data collected for measurement in EL6 on 2016.04.20.

4.1.7.4 Graph

The graph from the data in table 4.15 is shown in figure 4.15.

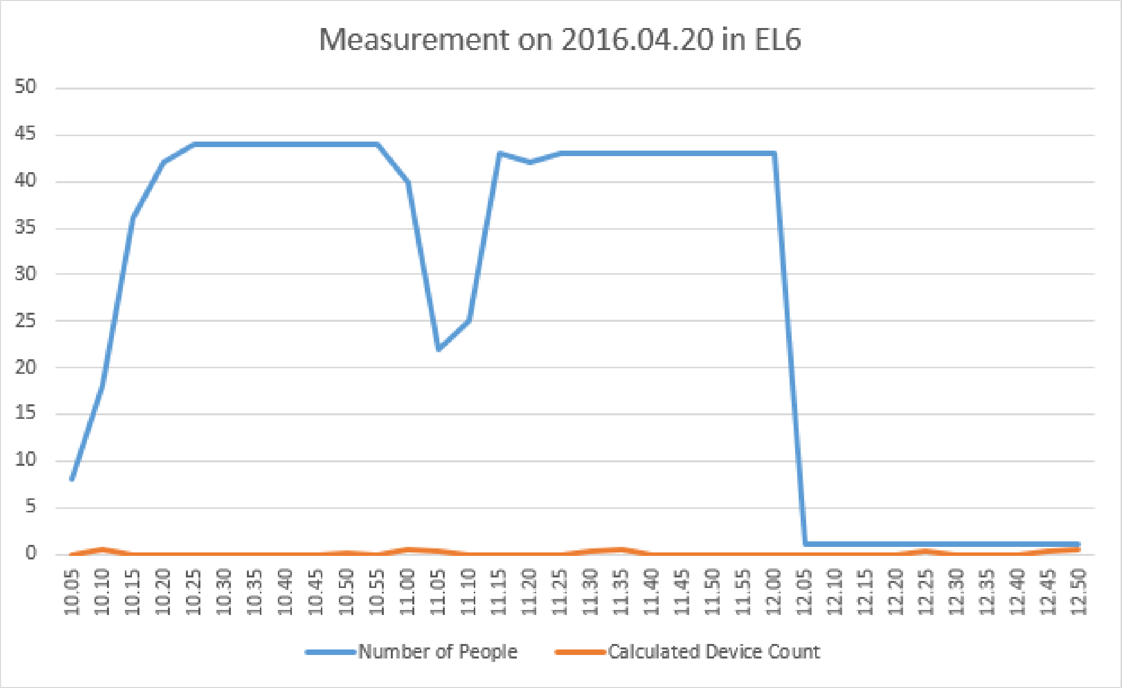


Figure 4.15: Graph visualizing the data from measurement on 2016.04.20 in EL6.

4.1.7.5 Correlation

The PPMCC from the data in table 4.15 is -0.16536.

4.1.7.6 Notes

- Notes by hand from blackboard.
- No laptops used.
- 1 AP in the room.

4.1.8 Measurement on 2016.04.20 in R8 and R9

4.1.8.1 Metadata

Information regarding the measurement is in table 4.16 below.

Room	R8 and R9
Building	Realfagsbygget
Capacity R8	150
Capacity R9	110
Date	2016.04.20
Time	13:00 to 14:00
Schedule R8	13:15 to 14:00: No lecture
Schedule R9	13:15 to 14:00: No lecture.

Table 4.16: Metadata for measurement in R8 (13:00 to 14:00) and R9 on 2016.04.20.

4.1.8.2 Map Location

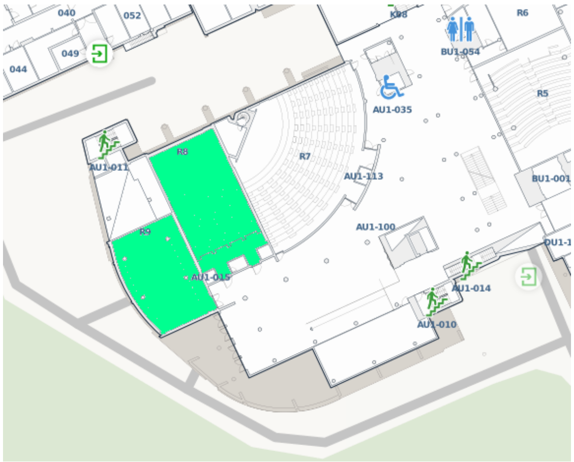


Figure 4.16: Map showing where H8 and R9 is located. The rooms are highlighted in green. R9 is in the bottom left and R8 in the top right.

4.1.8.3 Data

This measurement was conducted by viewing the doors to R8 and R9 from the hallway outside the lecture halls. It was possible to monitor both doors at the same time. An initial count was performed at 13:00, and the rooms were found to be empty. The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in two tables. Table 4.17 contains the data from room R8, and table 4.18 contains the data from room R9.

Time	**Number of People**	**Calculated Device Count**
13:00	0	0.0
13:05	0	0.0
13:10	0	0.0
13:15	0	0.0
13:20	0	0.0
13:25	0	0.0
13:30	0	0.640013
13:35	0	0.0
13:40	0	0.217291
13:45	0	0.0
13:50	2	0.0
13:55	2	0.245046
14:00	3	0.0

Table 4.17: Data collected for measurement in R8 on 2016.04.20 (13:00 to 14:00).

Time	**Number of People**	**Calculated Device Count**
13:00	0	0.0
13:05	0	0.0
13:10	0	0.0
13:15	0	0.0
13:20	0	0.0
13:25	0	0.0
13:30	0	0.0
13:35	0	0.0
13:40	0	0.0
13:45	0	0.0
13:50	1	0.0
13:55	1	0.0
14:00	1	0.0

Table 4.18: Data collected for measurement in R9 on 2016.04.20.

4.1.8.4 Graph

There are two graphs, one for each of the data tables. The graph from the data in table 4.17 is presented in figure 4.17, and the data in table 4.18 is presented in figure 4.18.

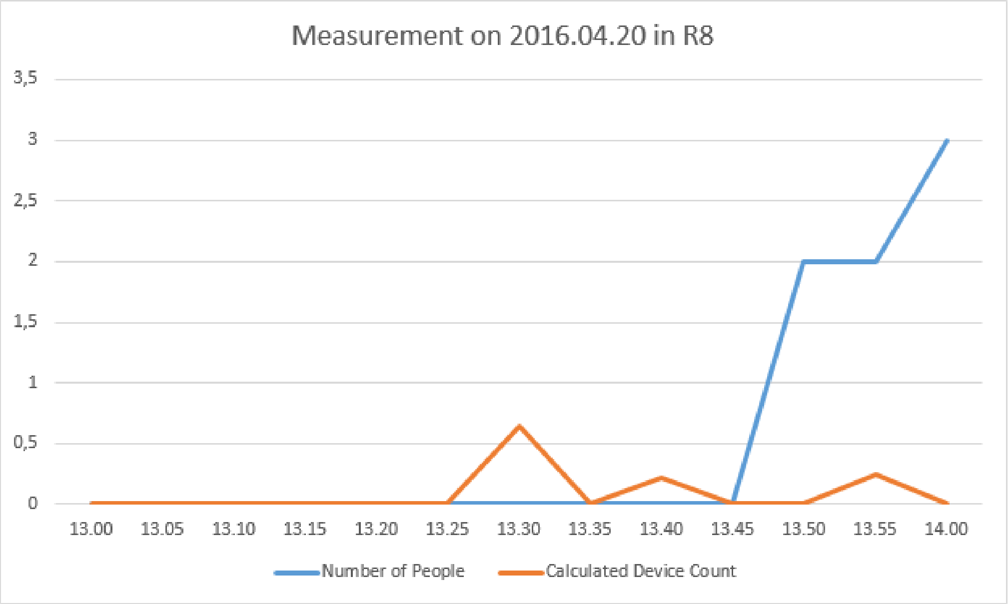


Figure 4.17: Graph visualizing the data from measurement on 2016.04.20 in R8 (13:00 to 14:00).

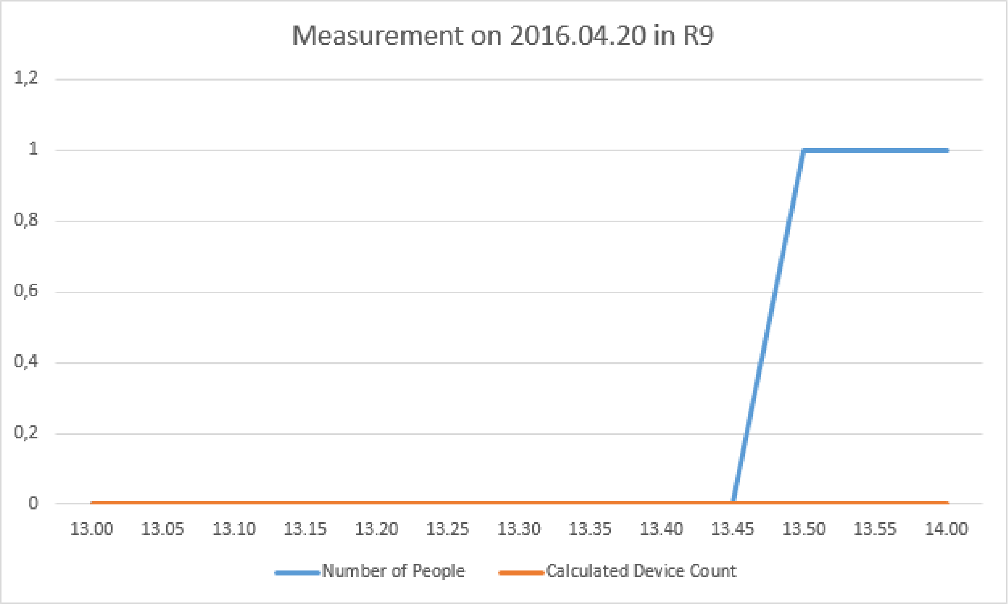


Figure 4.18: Graph visualizing the data from measurement on 2016.04.20 in R9.

4.1.8.5 Correlation

The PPMCC from the data in table 4.17 is -0.04372. It is not possible to calculate equivalent value for R9 as all *Calculated Device Count* values are zero. This causes division by zero in the formula.

4.1.8.6 Notes

- Measured both rooms from outside by looking at the only entrances for the rooms. The rooms were empty when the measurement first started.
- Suspects the result would've been more interesting if the rooms had been in use instead of being empty.

4.1.9 Measurement on 2016.04.20 in R8

4.1.9.1 Metadata

Information regarding the measurement is in table 4.19 below.

Room	R8
Building	Realfagsbygget
Capacity	150
Date	2016.04.20
Time	13:00 to 16:00
Schedule	13:15 to 14:00: No lecture
—”—	14:15 to 16:00: Lecture (TFY4230 – Statistisk Fysikk)

Table 4.19: Metadata for measurement in R8 on 2016.04.20 (13:00 to 16:00).

4.1.9.2 Map Location

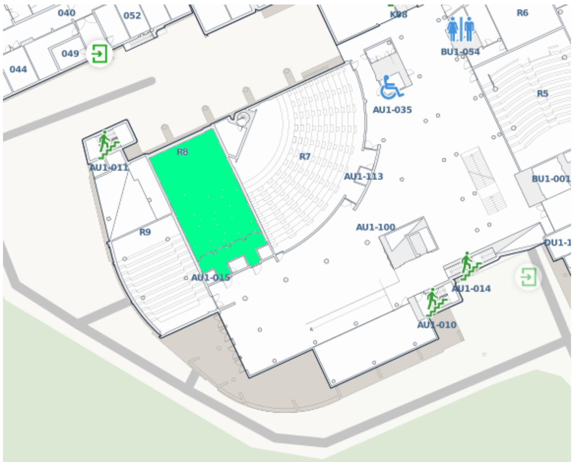


Figure 4.19: Map showing where R8 is located. The room is highlighted in green.

4.1.9.3 Data

The results from this measurement overlaps for one hour with the measurement in section 4.1.8 which was a measurement of both R8 and R9 at the same time. The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.20.

Time	**Number of People**	**Calculated Device Count**
13:00	0	0.0
13:05	0	0.0
13:10	0	0.0
13:15	0	0.0
13:20	0	0.0
13:25	0	0.0
13:30	0	0.640013
13:35	0	0.0
13:40	0	0.217291
13:45	0	0.0
13:50	2	0.0
13:55	2	0.245046
14:00	3	0.0
14:05	9	0.0
14:10	30	0.0
14:15	51	0.0
14:20	55	0.0
14:25	57	0.0
14:30	59	0.0
14:35	59	0.0
14:40	59	0.994904
14:45	59	0.0
14:50	60	0.0
14:55	60	0.0
15:00	60	0.994904
15:05	36	0.0
15:10	37	0.0
15:15	58	0.0
15:20	59	0.0
15:25	59	0.790792
15:30	59	0.683762
15:35	59	0.0
15:40	59	0.0
15:45	59	0.993599
15:50	59	0.994904
15:55	14	0.0
16:00	5	0.0

Table 4.20: Data collected for measurement in R8 on 2016.04.20 (13:00 to 16:00).

4.1.9.4 Graph

The graph from the data in table 4.20 is shown in figure 4.20.

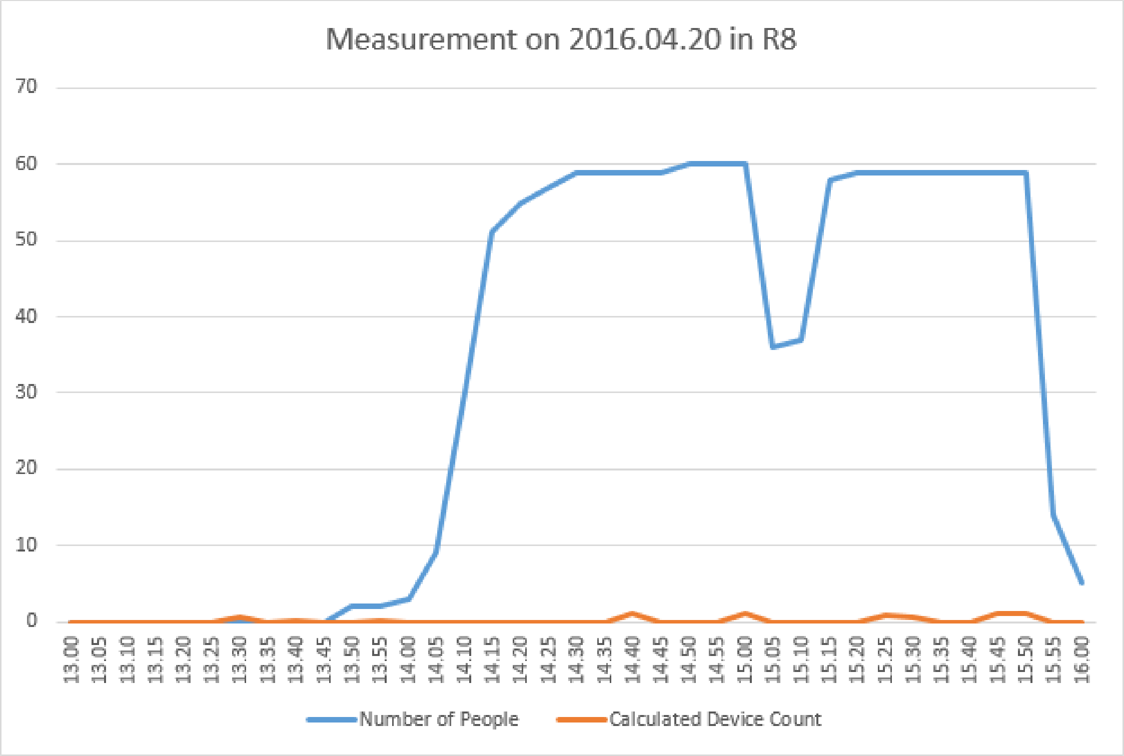


Figure 4.20: Graph visualizing the data from measurement on 2016.04.20 in R8 (13:00 to 16:00).

4.1.9.5 Correlation

The PPMCC from the data in table 4.20 is 0.326448.

4.1.9.6 Notes

- Notes by hand from blackboard.
- No laptops used.
- 1 AP in the room.

4.1.10 Measurement on 2016.04.21 in R9

4.1.10.1 Metadata

Information regarding the measurement is in table 4.21 below.

Room	R9
Building	Realfagsbygget
Capacity	110
Date	2016.04.21
Time	10:00 to 11:00
Schedule	10:15 to 11:00: Lecture scheduled, but room was empty

Table 4.21: Metadata for measurement in R9 on 2016.04.21.

4.1.10.2 Map Location

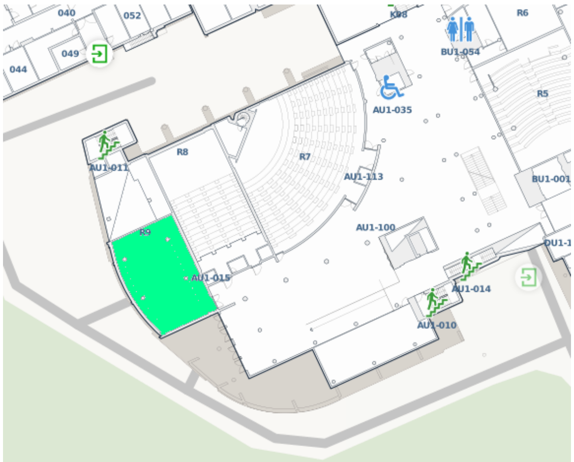


Figure 4.21: Map showing where R9 is located. The room is highlighted in green.

4.1.10.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.22.

Time	**Number of People**	**Calculated Device Count**
10:05	2	3.97961
10:10	1	0.0
10:15	1	0.0
10:20	1	0.0
10:25	1	0.0
10:30	1	0.0
10:35	1	0.0
10:40	1	0.0
10:45	1	0.0
10:50	1	0.0
10:55	1	0.0
11:00	1	0.0

Table 4.22: Data collected for measurement in R9 on 2016.04.21.

4.1.10.4 Graph

The graph from the data in table 4.22 is shown in figure 4.22.

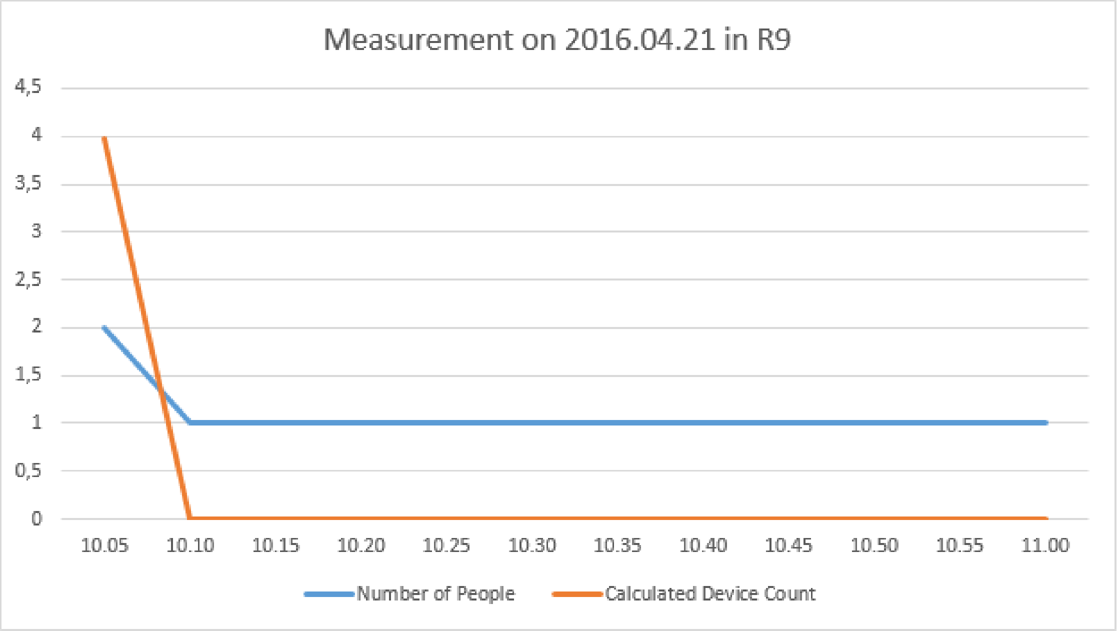


Figure 4.22: Graph visualizing the data from measurement on 2016.04.21 in R9.

4.1.10.5 Correlation

The PPMCC from the data in table 4.22 is 1.000.

4.1.10.6 Notes

- Cancelled lecture (BI2033 – Populasjonsøkologi).
- 1 AP in the room.

4.1.11 Measurement on 2016.04.22 in F2

4.1.11.1 Metadata

Information regarding the measurement is in table 4.23 below.

Room	F2
Building	Gamle Fysikk
Capacity	72
Date	2016.04.22
Time	10:00 to 12:00
Schedule	10:15 to 12:00: Lecture (TMA4275 – Levetidsanalyse)

Table 4.23: Metadata for measurement in F2 on 2016.04.22.

4.1.11.2 Map Location

Figure 4.23: Map showing where F2 is located. The room is highlighted in green.

4.1.11.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.24.

Time	**Number of People**	**Calculated Device Count**
10:00	5	0.0
10:05	5	0.0
10:10	8	0.994904
10:15	12	0.321957
10:20	16	0.0
10:25	16	0.0
10:30	16	0.994904
10:35	16	0.0
10:40	17	0.985162
10:45	17	0.377553
10:50	17	0.994904
10:55	17	0.0
11:00	17	0.994904
11:05	9	0.321957
11:10	16	0.0
11:15	16	0.0
11:20	17	0.640781
11:25	17	0.0
11:30	17	0.0
11:35	17	0.0
11:40	17	0.0
11:45	17	0.0
11:50	17	0.0
11:55	17	0.0
12:00	13	0.994904

Table 4.24: Data collected for measurement in F2 on 2016.04.22.

4.1.11.4 Graph

The graph from the data in table 4.24 is shown in figure 4.24.

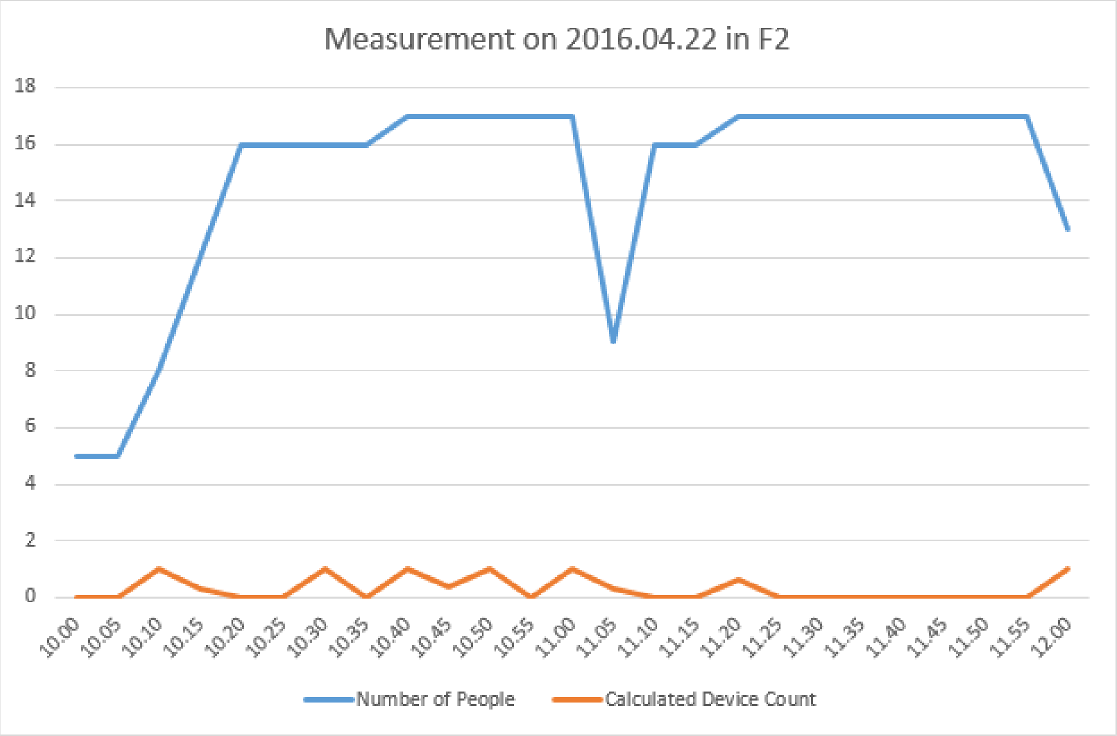


Figure 4.24: Graph visualizing the data from measurement on 2016.04.22 in F2.

4.1.11.5 Correlation

The PPMCC from the data in table 4.24 is -0.02649.

4.1.11.6 Notes

- Notes by hand from blackboard and power point presentation.

4.1.12 Measurement on 2016.04.22 in F6

4.1.12.1 Metadata

Information regarding the measurement is in table 4.25 below.

Room	F6
Building	Gamle Fysikk
Capacity	88
Date	2016.04.22
Time	12:00 to 14:00
Schedule	10:15 to 12:00: Lecture (TMA4212 – Numerisk løsning . . .)

Table 4.25: Metadata for measurement in F6 on 2016.04.22.

4.1.12.2 Map Location

Figure 4.25: Map showing where F6 is located. The room is highlighted in green.

4.1.12.3 Data

The researcher did the measure with an time interval of 5 minutes between each time count was noted down. The data from the measurement is shown in table 4.26.

Time	Number of People	Calculated Device Count
12:05	3	0.0
12:10	10	0.0
12:15	15	0.0
12:20	17	0.464024
12:25	17	0.0
12:30	17	0.0
12:35	17	0.0
12:40	17	0.0
12:45	17	0.87556
12:50	17	0.994904
12:55	17	0.43778
13:00	16	0.0
13:05	17	0.0
13:10	17	0.0
13:15	17	0.0
13:20	17	0.0
13:25	17	0.352639
13:30	17	0.683762
13:35	17	0.994904
13:40	17	0.0
13:45	17	0.0
13:50	17	0.994904
13:55	17	0.0
14:00	0	0.0

Table 4.26: Data collected for measurement in F6 on 2016.04.22.

4.1.12.4 Graph

The graph from the data in table 4.26 is shown in figure 4.26.

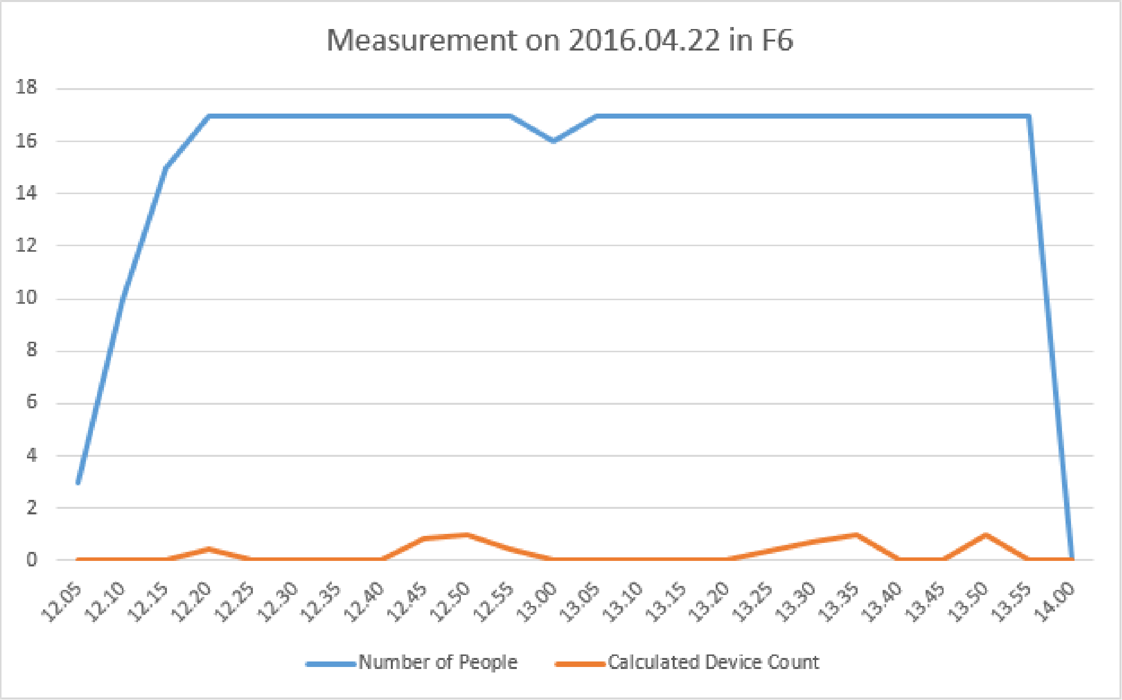


Figure 4.26: Graph visualizing the data from measurement on 2016.04.22 in F6.

4.1.12.5 Correlation

The PPMCC from the data in table 4.26 is 0.250516.

4.1.12.6 Notes

- Lecture: TMA4212 – Numerisk løsning av differanselikninger med differansemetoder
- Notes by hand from blackboard.
- During the break from 12:55 to 13:10, two laptops were in use as well as 3 to 5 smart phones.

4.2 Reading Rooms and Computer Labs

The measurements performed on computer labs and reading rooms are sorted by date and time. The earliest measurement is presented first.

4.2.1 Measurement on 2016.04.01 in Fiol

4.2.1.1 Metadata

Information regarding the measurement is in table 4.27 below.

Room	Fiol
Building	IT-bygget, sydfløy
Capacity	47
Date	2016.04.01
Time	14:15 to 16:15
Room type	Computer lab

Table 4.27: Metadata for measurement in Fiol on 2016.04.01.

4.2.1.2 Map Location

Figure 4.27: Map showing where Fiol is located. The room is highlighted in green.

4.2.1.3 Data

The researcher did the measurement with a time interval of 15 minutes between each time count was noted down. A full room count was performed at each interval. The reasoning behind a time interval of 15 minutes is that during such a long interval it is unlikely that a lot of people have entered or leaved the room, something which is more relevant when measuring lecture halls during break times. The data from the measurement is shown in table 4.28.

Time	Number of People	Calculated Device Count
14:15	5	3.00267
14:30	3	1.46683
14:45	5	1.19035
15:00	5	2.5538
15:15	6	1.43918
15:30	6	0.415987
15:45	2	0.705481
16:00	3	1.0701
16:15	2	1.35754

Table 4.28: Data collected for measurement in Fiol on 2016.04.01.

4.2.1.4 Graph

The graph from the data in table 4.28 is shown in figure 4.28.

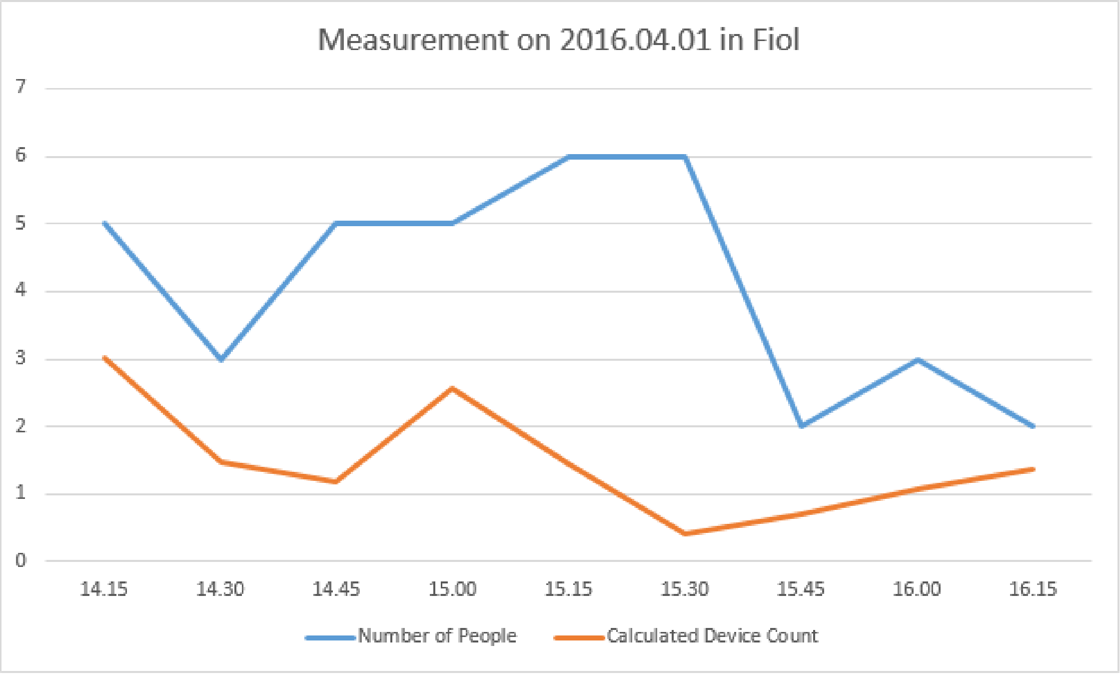


Figure 4.28: Graph visualizing the data from measurement on 2016.04.01 in Fiol.

4.2.1.5 Correlation

The PPMCC from the data in table 4.28 is 0.217706.

4.2.1.6 Notes

- A computer lab primarily for master students at IDI.
- Most occupants use personal laptops, but a few have borrowed a desktop computer from the department.
- 1 AP in the room. Placed close to the middle of the room on the roof.
- The large nearby lecture hall F1 was empty during the measurement.

4.2.2 Measurement on 2016.04.04 in Fiol

4.2.2.1 Metadata

Information regarding the measurement is in table 4.29 below.

Room	Fiol
Building	IT-bygget, sydfløy
Capacity	47
Date	2016.04.04
Time	13:30 to 15:30
Room type	Computer lab

Table 4.29: Metadata for measurement in Fiol on 2016.04.04.

4.2.2.2 Map Location

Figure 4.29: Map showing where Fiol is located. The room is highlighted in green.

4.2.2.3 Data

The researcher did the measurement with a time interval of 15 minutes between each time count was noted down. A full room count was performed at each interval. The data from the measurement is shown in table 4.30.

Time	Number of People	Calculated Device Count
13:30	18	12.174
13:45	18	13.2863
14:00	19	16.0989
14:15	10	2.84667
14:30	9	4.85863
14:45	10	3.2381
15:00	10	3.62158
15:15	13	4.65944
15:30	11	4.51061

Table 4.30: Data collected for measurement in Fiol on 2016.04.04.

4.2.2.4 Graph

The graph from the data in table 4.30 is shown in figure 4.30.

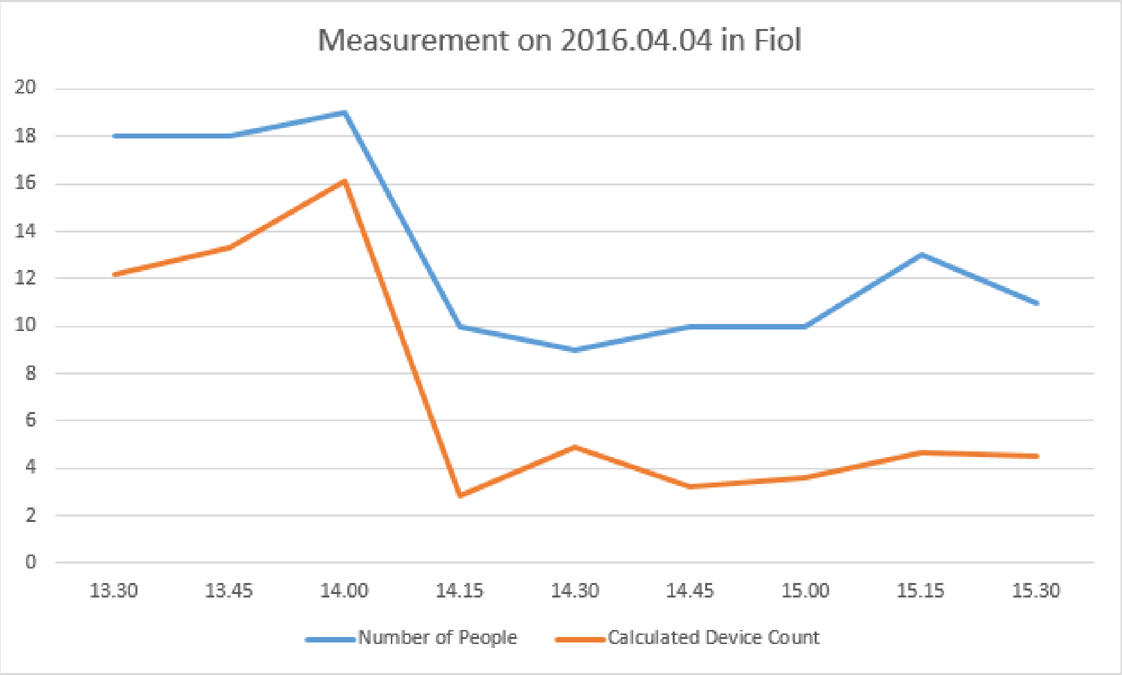


Figure 4.30: Graph visualizing the data from measurement on 2016.04.04 in Fiol.

4.2.2.5 Correlation

The PPMCC from the data in table 4.30 is 0.957743.

4.2.2.6 Notes

- As opposed to the previous measurement in Fiol, this time the lecture hall F1 is in use. There was a lecture from 12:15 to 14:00. The researcher checked the lecture hall in between time intervals to check if the room was in use or not. There was activity in the room until 14:00, but as of 14:05 the room was empty.

4.2.3 Measurement on 2016.04.11 in Fiol

4.2.3.1 Metadata

Information regarding the measurement is in table 4.31 below.

Room	Fiol
Building	IT-bygget, sydfløy
Capacity	47
Date	2016.04.11
Time	13:00 to 17:00
Room type	Computer lab

Table 4.31: Metadata for measurement in Fiol on 2016.04.11.

4.2.3.2 Map Location

Figure 4.31: Map showing where Fiol is located. The room is highlighted in green.

4.2.3.3 Data

The researcher did the measurement with a time interval of 15 minutes between each time count was noted down. A full room count was performed at each interval. The data from the measurement is shown in table 4.32.

Time	**Number of People**	**Calculated Device Count**
13:00	17	0.712441
13:15	18	0
13:30	18	1.3257
13:45	20	2.41632
14:00	17	6.44283
14:15	16	0.994904
14:30	20	0.20043
14:45	20	0.450264
15:00	16	2.22223
15:15	13	0.390401
15:30	12	1.95985
15:45	12	1.54652
16:00	12	1.73825
16:15	14	0.437185
16:30	12	0.870911
16:45	13	1.54736
17:00	13	2.43071

Table 4.32: Data collected for measurement in Fiol on 2016.04.11.

4.2.3.4 Graph

The graph from the data in table 4.32 is shown in figure 4.32.

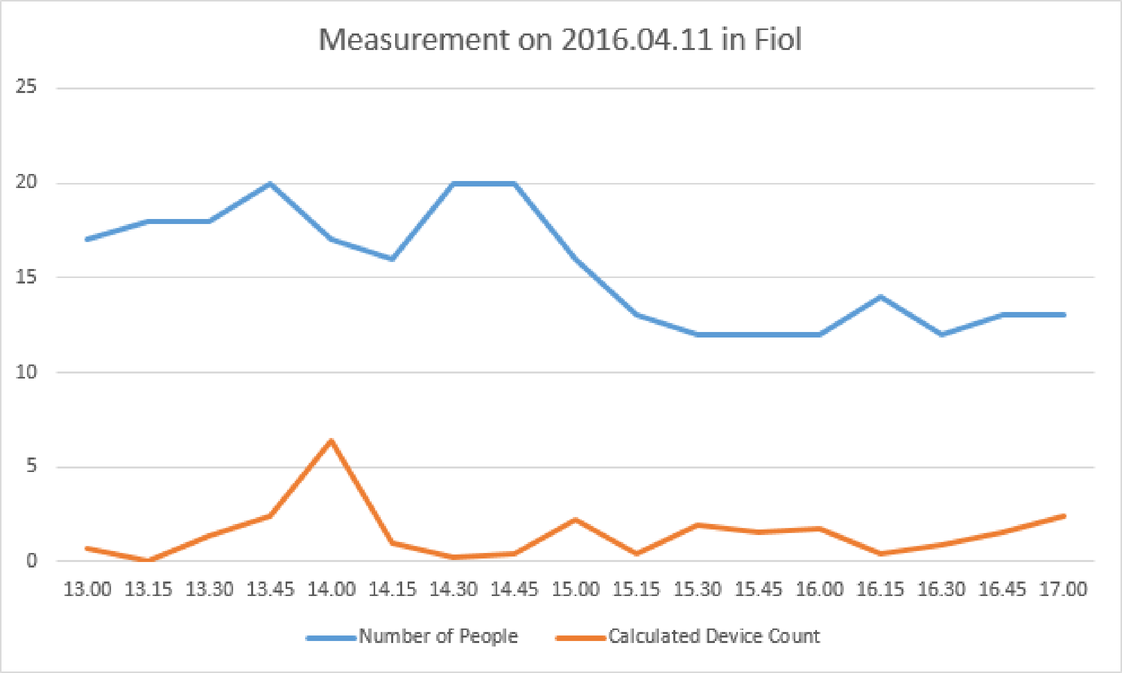


Figure 4.32: Graph visualizing the data from measurement on 2016.04.11 in Fiol.

4.2.3.5 Correlation

The PPMCC from the data in table 4.32 is −0.03809.

4.2.3.6 Notes

- There was a lecture in F1 during the measurement. The lecture was scheduled and in use from 14:15 to 16:00. Afterwards the room was empty until the measurement was finished (at 17:00).

4.2.4 Measurement on 2016.04.25 in Gribb

4.2.4.1 Metadata

Information regarding the measurement is in table 4.33 below.

Room	Gribb
Building	IT-bygget (basement floor)
Capacity	24
Date	2016.04.25
Time	10:00 to 12:00
Room type	Computer lab

Table 4.33: Metadata for measurement in Gribb on 2016.04.25.

4.2.4.2 Map Location

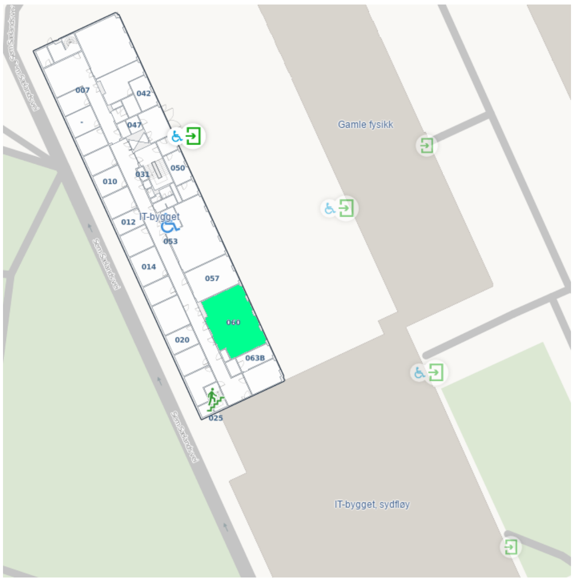


Figure 4.33: Map showing where Gribb is located. The room is highlighted in green.

4.2.4.3 Data

The researcher did the measurement with a time interval of 5 minutes between each time count was noted down. The counting strategy was to monitor the only door of the room and track the changes as people entered or left the room. A full room count was performed when the researcher first entered the room. The data from the measurement is shown in table 4.34.

Time	**Number of People**	**Calculated Device Count**
10:00	3	0.474261
10:05	5	0.230842
10:10	5	0.0
10:15	5	0.520032
10:20	5	0.0
10:25	5	0.0
10:30	6	0.0
10:35	6	0.0
10:40	7	0.0
10:45	7	0.0879283
10:50	6	0.0
10:55	8	0.0
11:00	8	0.0
11:05	8	0.0
11:10	8	0.0
11:15	9	0.0
11:20	8	0.0
11:25	9	0.0
11:30	9	0.0
11:35	8	0.0
11:40	9	0.0
11:45	9	0.0
11:50	10	0.0
11:55	10	0.0

Table 4.34: Data collected for measurement in Gribb on 2016.04.25.

4.2.4.4 Graph

The graph from the data in table 4.34 is shown in figure 4.34.

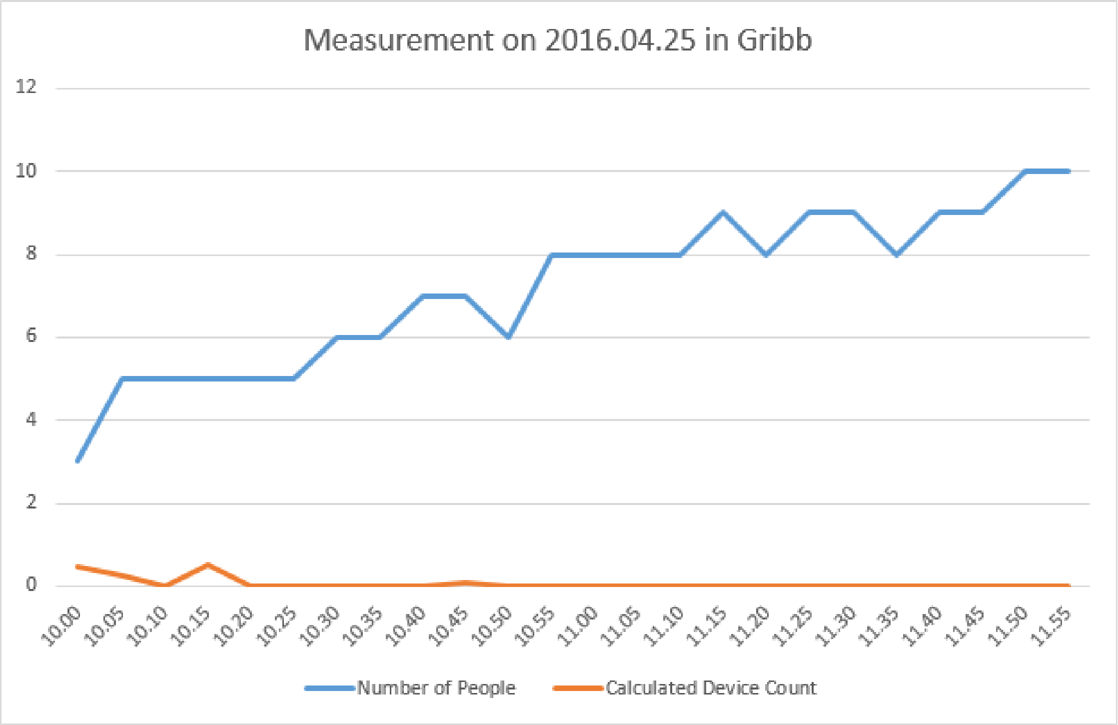


Figure 4.34: Graph visualizing the data from measurement on 2016.04.25 in Gribb.

4.2.4.5 Correlation

The PPMCC from the data in table 4.34 is -0.58275.

4.2.4.6 Notes

- The researcher was not able to locate an AP in the room. It could have been above the roof covers. The Wi-Fi signal strength on the researchers smart phone showed all bars were full, indicating excellent signal strength in the room.

4.2.5 Measurement on 2016.04.25 in Sule

4.2.5.1 Metadata

Information regarding the measurement is in table 4.35 below.

Room	Sule
Building	IT-bygget (2nd floor)
Capacity	20
Date	2016.04.25
Time	12:00 to 14:00
Room type	Computer lab

Table 4.35: Metadata for measurement in Sule on 2016.04.25.

4.2.5.2 Map Location

Figure 4.35: Map showing where Sule is located. The room is highlighted in green.

4.2.5.3 Data

Note that there are no numbers from 13:00 and onwards. The Analytics System was running but any queries performed on the API resulted in *No ticks found*. This is because of some error that occurred on the server the Analytics System runs on. See section 4.3 for a description of the error.

The researcher did the measurement with a time interval of 5 minutes between each time count was noted down. The counting strategy was to monitor the only door of the room and track the changes as people entered or left. A full room count was performed when the researcher first entered the room. The data from the measurement is shown in table 4.36.

Time	**Number of People**	**Calculated Device Count**
12:00	11	0.0
12:05	5	0.0
12:10	5	0.0
12:15	5	0.0
12:20	4	0.0
12:25	4	0.0
12:30	5	0.0
12:35	5	0.0
12:40	5	0.0
12:45	11	0.0
12:50	13	0.0
12:55	14	0.0
13:00	13	No data
13:05	13	No data
13:10	14	No data
13:15	13	No data
13:20	13	No data
13:25	13	No data
13:30	14	No data
13:35	13	No data
13:40	14	No data
13:45	13	No data
13:50	13	No data
13:55	14	No data

Table 4.36: Data collected for measurement in Sule on 2016.04.25.

4.2.5.4 Graph

The graph from the data in table 4.36 is shown in figure 4.36. The orange line stops at 12:55 as there are no data available from 13:00 and onwards.

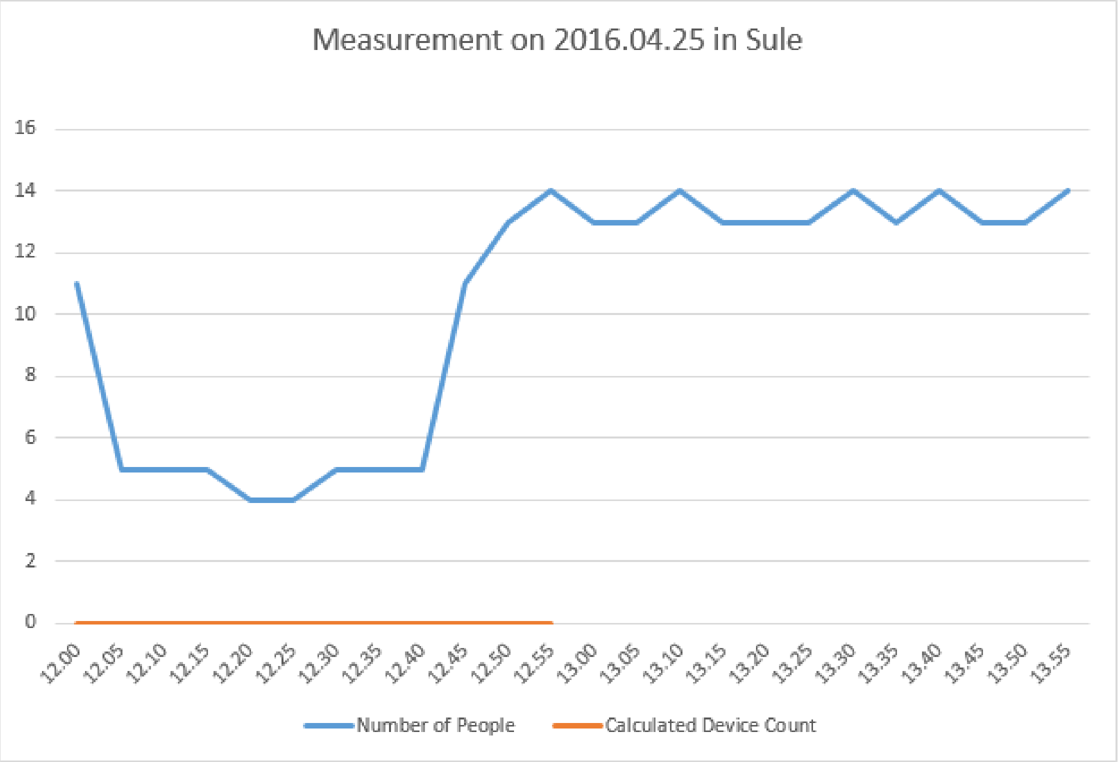


Figure 4.36: Graph visualizing the data from measurement on 2016.04.25 in Sule.

4.2.5.5 Correlation

The PPMCC from the data in table 4.36 was not possible to compute as it is not possible to divide by zero.

4.2.5.6 Notes

- Similarily to the previous measurement in Gribb, the researcher was not able to locate an AP in the room. It could have been above the roof covers. The Wi-Fi signal strength on the researchers smart phone showed all bars was full, indicating excellent signal strength in the room.

4.3 Other Measurements and Lack of Data

This section discussed the problem with the Analytics System which occurred during the research. The communication between the researcher and Wireless Trondheim where it was attempted to discover the reason for the problem happening is also presented.

4.3.0.1 Analytics System and Server Problems

The researcher had planned (and conducted) more measurements after the last conducted measurement in Sule on 2016.04.25. There was an error in the Analytics System which occurred at 13:00 2016.04.25. This has resulted in no more data is produced by the system after the time the error occurred. This means that all previous calculated device count data are still available, but new data is not produced. Therefore, the measurements which were planned is not presented as they do not contain data about device count, only how many people were in the rooms.

4.3.0.2 Problem Inquiry 1

After the researcher had contacted the man responsible for the Analytics System at Wireless Trondheim, Håkon Lorentzen, about the problem with the system, the response was that there had been some server problems and this had led to problems with the Analytics System. After a further inquiry by the researcher a few days later about the status of the problem, the response was that Wireless Trondheim have some raw data (the positioning data collected by the Cisco MSE) collected up to and including the 2nd of May, but the processing stopped sometime during the 9th of March 2016. There had been an attempt to kickstart the Analytics System by processing the remaining raw data, but it was unknown how long this would take. This response was on the 9th of May, but it appears that no new data has been produced by the Analytics System since the time the error occurred on the 25th of April 2016 at 13:00.

4.3.0.3 Problem Inquiry 2

An update mail received from Håkon Lorentzen on the 1st of June 2016 describes more details of what the problem was. The mail sent to Lorentzen by the researcher prior to receiving the mail was asking a few questions such as: ”Have anything changed since the Analytics System had been created?”, ”What was the reason for lack of positioning data?” and ”Had the Analytics System been used since it was created?”. The response stated that the system had not been used much at all since it was created by its original developer. It had just been running on a server and been processing data. The reply also stated the system is not greatly robust if other tasks or processes on the server requires more processing power. The processing logic by the Analytics System looks at the clock and processes data from -20 to -10 minutes (minus in front of the numbers means time relative to the current time). This means that if other tasks or processes such as backup of collected positioning data is being executed on the server, some data will never be processed by the

Analytics System. Lorentzen also stated that it was still unknown what caused the stop of processing of positioning data from the 9th of March 2016 and onwards. In the period from the 9th of March to the 25th of April there is no data, but from the 25th of April Lorentzen has manually processed some data.

Chapter 5

Discussion

The research has investigated whether or not the collected positioning data can be used to give an accurate representation of actual room usage on campus Gløshaugen. The research has been led by two major research questions (see section 3.1) which investigated the accuracy of the calculated device count data produced by the Analytics System, and how the accuracy can be improved or better handled. By using a quantitative research approach for conducting the research it has been possible to answer the research questions and therefore achieve the goal of the research.

The structure of the chapter is that it starts by answering the research questions and deciding whether or not the goal of the research was achieved. Afterwards both implications and limitations of the research is discussed.

5.1 Goal and Research Questions

In this section the results of the research is discussed in order to answer the research questions from section 3.1. Implications this research has on the research field of indoor positioning systems is presented, as well as limitations of the execution of the research. The section is structured such that each research question is presented in the same order as in chapter 3 for Research Design and Methodology. After all the questions have been attempted to be answered, the researcher decided whether or not the goal of the research is reached.

5.1.1 Room Representation using Calculated Device Count

> **RQ1**: How good representation of actual room usage does the calculated device count from the Analytics System provide?

The results show that in all cases the calculated device count is always lower than the actual room usage in all rooms. This applies for all the room types investigated: Lecture halls, reading rooms and computer labs. In some cases the calculated device count is poor to such a degree that it is impossible to derive actual room usage from calculated device count. In these cases, the numbers of calculated device count is barely above zero throughout the measurements. For instance, the measurement in EL5 on 2016.04.18 (see section 4.1.5), there are periods of time where there are about 150 people present in the room but the calculated device count is a number which fluctuates between 0 and 10, but mostly stays below 5. The measurement in R7 on 2016.04.13 (see section 4.1.1) had a similar number of people in the room: About 155 people. The numbers produced by the Analytics System ranged from 2 to 28, and most often the number was 10 or larger. It is weird that lecture halls with about 150 people in them produces such a low calculated device count by the Analytics System. Measurements in S3 on 2016.04.15 (see section 4.1.3) also had about 150 people, but the results showed that the calculated device count was poor in this measurement too with a number averaging between 0 and 3.

For rooms where the measured number of people were lower than in the large rooms the calculated device count were also poor. What is interesting to note is that there usually is not a large difference between rooms which was nearly empty and rooms with some people in them. The measurement in H3 on 2016.04.20 shows that there was no significant difference in the calculated device count when there was a lecture with about 40 attendees and when the room was empty (except for the researcher) after the lecture was over. A similar result was found in the measurement in R8 on 2016.04.20 (see section 4.1.9). There were about 50 to 60 people present during the lecture from 14:00 to 16:00. The measurement started at 13:00 when the room was empty. The calculated device count is very similar for when the room was empty and when in use. The same can be said of the measurement in El6 on 2016.04.20 (see section 4.1.7) where there at most was 44 people present but the calculated device count didn't change much when the room was in use compared to when it was empty.

There are few options at Gløshaugen where a single researcher can do a measurement of two rooms at the same time as it is required that the researcher is able to monitor the entrances for two rooms at the same time while keeping track of how many people enters or leaves the rooms. A measurement where the researcher was able to measure two rooms at once was for the rooms R8 and R9 on 2016.04.20. (see section 4.1.8). It was possible for the researcher to to the measurement from the hallway outside of the rooms as it was viable to have line-of-sight to both entrance doors of the rooms at the same time in order to keep track of the room count. An initial room count was performed when the researcher started the measurement, and it was found that the rooms were empty at this time. The original idea of doing a measurement of two rooms at the same time is that it could possible to see if a spillage effect has happened. This is when devices in room A is found out to be in a nearby room B as there are some inaccuracies related to the positioning data collected by the Cisco MSE (see section 2.2.2). A spillage effect would result in the calculated device count graph being similar for both rooms, but since the rooms unfortunately were empty for most of the time during the measurement no such spillage effect occurred. Although the rooms were empty, the canteen on the floor right above the rooms was in use. This is the canteen of Realfagsbygget. It is not known exactly how many people were using the canteen at the time of the measurement, but a guess by the researcher suspects that there are always at least 20 people there during the day. But there will often be more people. The results showed that there was no spillage effect where devices from the floor above the lecture halls were found to be in the lecture halls. This could either suggest that the accuracy in the z-dimension for the positioning data is better than in the xy-plane, but with the poor results of calculated device counts in general from all the measurements it is not possible to draw any conclusion from this.

The result discussed in the above three paragraphs makes it evident that the calculated device count data from the Analytics System is not suited for use for representing actual room usage at campus Gløshaugen. Therefore RQ1 has been answered, albeit not in the way the researcher had hoped when starting the work in this research project.

5.1.1.1 Room Type and Room Representation

> **RQ1.1**: How does the representation of room usage change depending on the type of room being used?

There were also poor results in the measurements done for reading rooms and computer labs. These types of rooms are often much smaller than lecture halls. For instance, Fiol has a capacity of 47 persons. Fiol was also the largest computer lab which was measured. In the measurement in Fiol on 2016.04.11 (see section 4.2.3) there were around 15 people in the room at all times (12 at lowest and 20 at most). The calculated device count usually fluctuated between 0 and 2 with occasional spikes up to 6. With such low calculated device count data it is difficult to give an actual representation of room usage. The calculated device count must in this case be multiplied by 10 in order to get usage data numbers to be close to the real number of people in the room. Similar results were found in the other measurements in Fiol (see section 4.2.1 and 4.2.2) as well as Gribb (see section 4.2.4).

Given that the results were poor for reading rooms and computer labs as well as lecture halls it is not possible to determine if room type has an effect of how the calculated device count can be used to give a good representation of room usage in a room. This answers RQ1.1 as there is no apparent change in the calculated device count depending on the room type.

5.1.1.2 Linear Relationship Between Number of People and Calculated Device Count

> **RQ1.2**: What is the linear relationship between the calculated device count and the number of people in a room?

Since the numbers produced by the Analytics System has generally been low, the Pearson Product-Moment Correlation Coefficient (PPMCC, see section 3.5) values calculated for all measurements has not made it possible to determine if there actually is a linear relationship between the number of people in a room and the calculated device count from the Analytics System. At best it is possible to determine if a lecture hall has been in use or not. But this requires that there has been a sufficient number of people in the room, such as the measurements where there was about 150 people present. In some cases it has not been possible to calculate a PPMCC value as the Analytics System has reported 0.0 usage values during the measurements. In other cases the values has been varying, but usually not higher that 0.3, and sometimes even negative. In one measurement in Fiol on 2016.04.04 a PPMCC value of 0.957743 was achieved. The reason for this is believed to be a coincidence as the nearby large lecture hall F1 was being emptied of people after a lecture at the same time nearly half the people in Fiol left the room. This does suggest that there is a spillage effect due to the inaccuracies of the positioning data as it has been likely that position data from devices in F1 has been found to be in Fiol instead.

Because no linear relationship can be drawn from the results, the answer to RQ1.2 is that there is no apparent linear relationship between the calculated device count and the number of people.

5.1.2 Factors Affecting the Accuracy

RQ2: How can the accuracy be improved or the inaccuracy be better handled?

Since the results were poor in general for all measurements it has been difficult to determine what factors affect the accuracy of the positioning data. It was apparent that there are certain areas at Gløshaugen where the accuracy is poorer than other places. For example the entrance to IT-bygget and the 13th floor in Sentralbygg 1 (see section 3.3.1). Due to limitations with the Analytics API (discussed later in section 5.3.2), some rooms have not been available for the researcher. It was discovered by both Aulie [2015] and Eriksen [2015] that the accuracy in F1 and the nearby rooms were poor. If F1 had been available in the Analytics API the researcher would have preferred to conduct measurements in the room in order to investigate the accuracy in the area.

During the research some APs on the first floor of IT-bygget were upgraded with the Cisco Hyperlocation module. Some of the APs nearby, but not inside, Fiol were upgraded. The researcher wished to conduct a measurement in Fiol after the upgrade took place in order to compare the result with how the calculated device count was before the modules were upgraded. But since the Analytics System stopped producing data right before the modules were put in place it was not possible to to this before-and-after comparison of the Hyperlocation module.

Given the results and no discovered apparent factors which affected the accuracy of the positioning data, it is not possible to answer RQ2 in this research.

5.1.3 Research Goal

> **Goal**: Investigate if the number of devices calculated from collected positioning data can give an accurate representation of room usage in rooms at campus Gløshaugen.

As previously discussed it was possible to answer RQ1, RQ1.1 and RQ1.2. These are the questions related to whether or not the calculated device count from the Analytics system can be used for representing room usage at Gløshaugen. It was not possible to draw any answers to RQ2 which attempted to determine factors which affected the accuracy of the positioning data, or how the inaccuracy could be better handled.

Even though it was not possible to answer RQ2, all the other questions were answered. This means the goal of the research has been achieved as it is reasonably apparent that the data produced by the Analytics System can not be used to give an accurate representation of room usage at all.

5.2 Implications

The research has investigated if the collected positioning data could be used to give an accurate representation of room usage at campus Gløshaugen. In this section, the implications the results from this project have on the research field and future work is discussed.

5.2.1 Usefulness for Room-Representation

The results from the research suggest that the data can not be used to accurately give a good representation of room usage on campus Gløshaugen. As Håkon Lorentzen stated in a mail regarding the Analytics System's lack of processing of positioning data (see section 4.3), the system is not robust as it will not process data at all because it yields processing power if there are other processes executing on the same server which the Analytics System runs on. It is strongly suggested to make the current system robust in order to be able to process all the gathered positioning data.

5.2.2 Previous Research

The results from this research suggests that other applications that have used the data produced by the Analytics System produces inaccurate data, especially since it has become clear that the Analytics System has not changed since it was created (see section 4.3). Previous projects, such as Thingstad and Tran [2014] and Mersland [2013] did rely on the data from the Analytics System in the prototype applications which had been developed in their projects. Thingstad and Tran [2014] created a prototype mobile application which helps students in finding reading rooms which has available spots, while Mersland [2013] developed a prototype web application which could be used for facility managers for determining room usage for various rooms at Gløshaugen.

Other research projects, such as Eriksen [2015] and Aulie [2015], which visualized human mobility patterns using heatmaps, used the raw positioning data collected by the Cisco MSE and not the usage data produced by the Analytics System. This means that the results from this research project can not conclude that data produced by applications which relies on the raw data has inaccuracies in them. As the calculated device count were generally poor in all the measurements conducted, it is neither possible to say for sure that some locations at Gløshaugen has better accuracy of the positioning data than other locations.

5.2.3 Future Work

This research project and its results suggests that the usage data produced by the Analytics System can not be used to give an accurate representation of room usage in its current state. The researcher would first advice that the problems with the system regarding not being able to process all the positioning data (see section 4.3) should be fixed. This problem can be solved by having the Analytics System run in its own dedicated server which would ensure that the processing would not yield to other tasks and processes running on the same server.

Another suggestion is that the processing logic in the algorithm which calculates device count usage is rewritten. The reason is that the current source code does not explain what mathematical and statistical methods are used when determining what room a device is located in as well as determining the spillage effect a device has on nearby rooms. For instance, the current algorithm, which uses a method of splitting a device to nearby rooms depending on the accuracy associated with the positioning data of the device, does not explain through documentation what method is being used to calculate the spreading out of devices. Some magic numbers does also occur in the source code. This makes it difficult for anyone working with the Analytics System to know for sure how the algorithm affects the resulting produced usage data.

If Wireless Trondheim chooses to improve the system either by giving it more processing power or rewriting the source code, a future project could be performed where similar measurements as has been conducted in this research project is done in order to determine if the accuracy of the usage data is improved or not. For instance, a group from Experts in Team (a subject at NTNU which focuses on group work with people with various background) could do the measurements and assess the accuracy of the improved system. When a group consists of 5 to 6 people, measurements can be conducted at the same time. This makes it possible to measure rooms in parallel more accurately than a single researcher is able to. There are many types of Experts in Team classes with different focuses, but a class with focus on *Big Data* would be more fitting for this type of project.

It would also be interesting to do more work on determining how many APs are nearby a room in order to see how the number of APs affects the accuracy of the usage data. Additionally, it would be interesting to do before and after measurements of rooms where nearby APs have been upgraded with the Cisco Hyperlocation Module (see section 2.2.2.2) to see if the accuracy increase from these modules would greatly improve the positioning data. It is also curious to know if the orientation of APs affects the accuracy. The Cisco APs is meant to be mounted horizontally in order to give the best accuracy for locating the position of a device. In some rooms the APs must be mounted vertically as they are placed on the wall instead of the roof.

5.3 Limitations

This section provides limitations which has affected the outcome of the research.

5.3.1 Analytics System Performance

As was discovered in an email exchange between the researcher and the man currently responsible for the Analytics System at Wireless Trondheim, it was found that the present Analytics System is not robust as it would need to yield to other processes running on the same server. Examples of such processes is doing backup of all positioning data. These processes are performed twice a day at around noon and midnight. As a result of the Analytics System yielding, not all of the collected positioning data will be processed as the algorithm only works with data which was collected from 10 to 20 minutes ago. If enough time passes and the algorithm hasn't been able to process data for a given time period these data will not be processed at all unless they are manually submitted by a developer. The goal of the research was to investigate if the calculated usage data could be used to give a good representation of actual room usage, and through this research it was determined that the usage data was inaccurate to such a degree it would not be wise to use it for applications which determines how many people are in a room. This is especially true as it became known during the research that the Analytics System is not able to process all positioning data at all times.

5.3.2 Analytics API Room Limitations

There were some limitations in room selection for measurements. The reason is that some rooms does not exist in the Analytics API. Examples of rooms that is not in the system is the large lecture hall F1, the reading rooms B1-165 and Drivhuset (located in IT-bygget sydfløy) and many rooms in Kjel-bygget. All of there are rooms the researcher had been interesting in doing measurements in. For instance, no measurements had been performed in Kjel-bygget. If this would have been possible the measurements would cover more buildings of campus Gløshaugen. F1 would also be interesting to measure as it is perhaps the largest lecture hall on the campus. The lectures in F1 are usually the subjects which is taken by student from various departments and study programs.

5.3.3 Data Collection

Being a single researcher has made it difficult to do measurements in all the rooms at Gløshaugen. Due to time restraints the researcher had to choose a select few rooms while trying to distribute the measurements to cover different buildings and various room types. Being two or more researchers for collecting the actual room usage data would help in being able to cover more grounds faster. It would be possible for a master project with a single researcher to cooperate with a group of people for taking the measurements in different rooms. For instance, as mentioned in 5.2.3, an Expert in Team group which focuses on *Big Data* would be fitting as comparing large amounts of data of actual room usage with the calculated device count from the Analytics System would fall under the big data category.

An alternate way to determining how many people are in a room is to install automatic counters at doors. Installing a counter which is able to determine if a person enters or leaves a room would make it possible to do measurements over a longer period of time while and requires a minimal amount of time by a researcher. Setting up an automatic people counter system is expensive at is would require advanced hardware capable of accurately detecting how many people enters or leaves a room. Collecting the room usage data would also require a centralized server which is able to collect the data from an automatic counter. Setting up a server and the infrastructure for communication between the server to one ore more person counters would also increase the setup cost. Some rooms does also have more than a single entrance. This would require the server to hold an internal state of all person counters which is connected to the same room and monitor each of them simultaneously in order to be able to determine the number of people in a room. The development and testing of such an automatic person counter system could be an interesting project to conduct at NTNU.

Chapter 6

Conclusion

6.1 Conclusion

The research has been guided by using the quantitative research approach in order to study whether or not the collected positioning data can be used to give an accurate representation of room usage on campus Gløshaugen. By collecting data of how many people use various rooms at campus and comparing it to device usage data produced by the Analytics System, which produces its data based on the continuously collected positioning data from wireless Wi-Fi devices at campus, it was possible to compare the data to see if the accuracy of the positioning data could be used to give an accurate representation of room usage. Various factors such as room type and room size were investigated in order to learn if they affected the accuracy of the data produced by the Analytics System.

The research was steered by two major research questions which focused on accomplishing the goal of the research. The research questions are:

- **RQ1:** How good representation of actual room usage does the calculated device count provide?
  - **RQ1.1:** How does the representation of room usage change depending on the type of room being used?
  - **RQ1.2:** What is the linear relationship between the calculated device count and the number of people in a room?
- **RQ2:** How can the accuracy be improved or the inaccuracy be better handled?

The results from this research shows that the usage data produced by the Analytics System can't be used to give a good representation of actual room usage at Gløshaugen. In general, the numbers produced by the system are wither very low or none for most rooms measured. Only in rooms where there had been about 150 people present had it been possible to see a small spike in the usage data. Through communication with Wireless Trondheim it has become clear that the problem with the current system is that its processing must yield if there are other more important processes running on the same server the system is located. This results in the system not being able to process all of the collected positioning data as the algorithm only works with positioning data from a small time interval is processed. Once the time interval has been passed the data will no longer be processed as the system will now start working with positioning data from a newer time interval.

6.2 Future Work

The current version of the Analytics System needs to be improved in order to be able to produce better usage data. One option is to move the system to a dedicated server which would ensure that the systems processing doesn't have to yield to other processes on the server. This option is discussed more thoroughly in section 5.3.1.

Another thing which was discovered about the Analytics System is that the source code is undocumented. The original developer of the system is no longer at Wireless Trondheim. This makes it difficult to understand the inner workings of how the system calculates the device count for each room. Especially the method used for handling the inaccuracy of the collected positioning data, such as devices having a spillage effect to nearby rooms. The inaccuracy comes from MazeMap only being able to determine an object's position with an accuracy of up to 5-10 meters [Biczok et al., 2014].

If Wireless Trondheim were to improve the current Analytics System in the future, further work could be done to measure the accuracy of the produced usage data of the improved system to see if it has improved. Additionally, being more than a single researcher would make it possible to do measurements and therefore collect more data regarding room usage at Gløshaugen. It would be fitting for a group, such as a group from an Experts in Team class, to do measures as these groups often consists of 5 to 6 persons. This would make it possible to do measurements which covers more of Gløshaugen. This would help in determining if there are some areas on campus where the accuracy of the positioning data is better or worse.

Conducting before-and-after measurements of rooms where nearby access points have been upgraded with the Cisco Hyperlocation Module would determine if the accuracy of the positioning data improves when these module have been installed. Cisco boasts that the Hyperlocation module is capable of improving the localization accuracy of access points when determining an object's position to be from 1-3 meters (see section 2.2.2.2).

Bibliography

M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, and A. Hopper. Implementing a sentient computing system. *Computer*, 34(8):50–56, 2001.

S. H. Andresen, J. Krogstie, and T. Jelle. Lab and research activities at wireless trondheim. In *Proceedings of IEEE International Symposium on Wireless Communication Systems*, pages 385–389, 2007.

K. G. Aulie. Human Mobility Patterns from Indoor Positioning Systems. Master's thesis, Norwegian University of Science and Technology, 2015.

G. Biczok, S. Diez Martinez, T. Jelle, and J. Krogstie. Navigating mazemap: indoor human mobility, spatio-logical ties and future potential. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*, pages 266–271. IEEE, 2014.

Cambridge. The Bat Ultrasonic Location System. URL `http://www.cl.cam.ac.uk/research/dtg/attarchive/bat/`. Accessed on 2016-05-05.

Cisco. Cisco CleanAir Technology. URL `http://www.cisco.com/c/en/us/solutions/enterprise-networks/cleanair-technology/index.html`. Accessed on 2016-06-06.

Cisco. *Cisco Hyperlocation Module with Advanced Security.* Cisco Systems Inc, May 2016. Data sheet: C78-734901-02. URL: `http://www.cisco.com/c/en/us/products/collateral/interfaces-modules/aironet-hyperlocation-module-advanced-security/datasheet-c78-734901.pdf`.

J. Collin, O. Mezentsev, G. Lachapelle, et al. Indoor positioning system using accelerometry and high accuracy heading sensors. In *Proc. of ION GPS/GNSS 2003 Conference*, pages 9–12, 2003.

K. Curran, E. Furey, T. Lunney, J. Santos, D. Woods, and A. McCaughey. An evaluation of indoor location determination technologies. *Journal of Location Based Services*, 5(2):61–78, 2011.

L. Demir. *Wi-Fi tracking: what about privacy.* PhD thesis, M2 SCCI Security, Cryptologyand Coding of Information-UFR IMAG, 2013.

J. B. Eriksen. Visualization of Crowds from Indoor Positioning Data. Master's thesis, Norwegian University of Science and Technology, 2015.

N. Fallah, I. Apostolopoulos, K. Bekris, and E. Folmer. Indoor human navigation systems: A survey. *Interacting with Computers*, page iws010, 2013.

Y. Fukuju, M. Minami, H. Morikawa, and T. Aoyama. Dolphin: An autonomous indoor positioning system in ubiquitous computing environment. In *WSTFEUS*, pages 53–56, 2003.

Google. About - Google Maps. URL `http://www.google.com/maps/about/partners/indoormaps/`. Accessed on 2016-05-09.

Y. Gu, A. Lo, and I. Niemegeers. A survey of indoor positioning systems for wireless personal networks. *Communications Surveys & Tutorials, IEEE*, 11(1):13–32, 2009.

J. Hightower and G. Borriello. Location sensing techniques. *IEEE Computer*, 34 (8):57–66, 2001a.

J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, (8):57–66, 2001b.

B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle. *GNSS–global navigation satellite systems: GPS, GLONASS, Galileo, and more.* Springer Science & Business Media, 2007.

B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global positioning system: theory and practice.* Springer Science & Business Media, 2012.

J. J. Hox and H. R. Boeije. Data collection, primary vs. secondary. *Encyclopedia of social measurement*, 1:593–599, 2005.

K. Kaemarungsi and P. Krishnamurthy. Properties of indoor received signal strength for wlan location fingerprinting. In *Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004. The First Annual International Conference on*, pages 14–23. IEEE, 2004.

T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg. Compass: A probabilistic indoor positioning system based on 802.11 and digital compasses. In *Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization*, pages 34–40. ACM, 2006.

J. Little and B. O'Brien. A technical review of cisco's wi-fi-based location analytics, 2013.

H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1067–1080, 2007.

S. D. Martinez. Campusguiden: Indoor positioning, data analysis and novel insights. Master's thesis, Norwegian University of Science and Technology, 2013.

D. J. Mersland. Produktutvikling an nettaplikasjon for bruk of visualisering av posisjonsdata: Fra idé til prototype. Master's thesis, Norwegian University of Science and Technology, 2013.

D. Muijs. *Doing quantitative research in education with SPSS*. Sage, 2010.

K. Pahlavan. *Principles of wireless networks: A unified approach*. John Wiley & Sons, Inc., 2011.

S. Taneja, A. Akcamete, B. Akinci, J. Garrett, L. Soibelman, and E. W. East. Analysis of three indoor localization technologies to support facility management field activities. In *Proceedings of the International Conference on Computing in Civil and Building Engineering, Nottingham, UK*, 2010.

T. Thingstad and H. T. Tran. A location-based mobile service utilizing anonymous indoor user location data. Master's thesis, Norwegian University of Science and Technology, 2014.

UiO. Personvern og lokasjonsbaserte tjenester - Institutt for privatrett. URL `http://www.jus.uio.no/ifp/om/organisasjon/afin/forskning/notatserien/2001/1_01.html`. Accessed on 2016-05-06.

Wifarer. Wifarer • indoor positioning | indoor GPS | our story. URL `http://www.wifarer.com/our-story`. Accessed on 2016-05-11.

Wikipedia. Pearson product-moment Correlation Coefficient. URL `https://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient`. Accessed on 2016-06-01.

G. S. Wong. Speed of sound in standard air. *The Journal of the Acoustical Society of America*, 79(5):1359–1366, 1986.