NTNU

Norwegian University of
Science and Technology

Capability Assessment of Indoor Positioning Systems

Andreas Dypvik Landmark

Master of Science in Computer Science
Submission date: July 2009
Supervisor: Pieter Jelle Toussaint, IDI

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

Problem Description

Indoor Positioning Systems (IPS) technology is being used for more and more applications, ranging from inventory control to process monitoring and improvement. The health care sector has also showed interest in this rather novel technology and the emerging opportunities that can be realised through automatically deriving insight about the work environment from the use of sensor equipment. However, different IPS systems are based on a variety of methods for locating objects. These methods produce slightly different capabilities with inherent strengths and weaknesses which have to be taken into consideration when designing an installation. Hospitals wanting to implement an IPS to solve a specific need will not necessarily be aware of these challenges or possess the knowledge to make well-informed decisions.

The purpose of this Master's Thesis is to investigate how to assess the required capabilities of a location system and to investigate how these capability requirements can be assessed experimentally.

Assignment given: 15. January 2009
Supervisor: Pieter Jelle Toussaint, IDI

Abstract

Location systems are seen as a promising technology for tracking people and objects to improve efficiency and quality in the healthcare domain. To increase the chances of success when introducing this new technology there are certain operational capabilities that need to be understood. The purpose of this Thesis is to explore how these operational capabilities can be assessed by experiment.

The thesis proposes a method for describing the operational capabilities of a location system using a two-dimensional matrix of purposes of location systems in the healthcare domain, as found in literature. Using this matrix it is possible to assess and predict the requirements for a location system based on a classification of the purpose of the installation. Conversely it is possible to use the same matrix to find purposes that can be solved with a given location system.

Using the Sonitor Indoor Positioning System it was also demonstrated how the operational capabilities of a location system could be found through a series of small low cost and low effort experiments.

In conclusion three dimensions relating to operational capabilities were identified: granularity, resolution, and concurrency. Granularity and concurrency were shown to be successfully assessed through experiment, while resolution was found analytically. We also found a method to predict the impact of infrastructure size on the operational capability of the location system based on the same small experiments.

Preface

The work with this Thesis has been performed at Department of Computer and Information Science at NTNU as well as The Norwegian Electronic Health Record Research Centre (NSEP) during the spring semester of 2009.

The original work for this Thesis was intended to focus on the implementation of an IPS in a specific platform for supporting collaboration in a peri-operative environment. Due to usual circumstances, in which things usually fall in to place just a little too late, the focus of the work was changed to the more general issue of assessing the capabilities of a location system.

While undertaking this work I have been supervised by Assoc Prof. Pieter J. Toussaint which has offered excellent input and guidance throughout the process along with Post doc. Andreas R. Seim who also was of the utmost help and offered both tips and been a part of several fruitful discussions between the three of us. A further thank you is in order for the rest of the participants of the COSTT-project (Arild, Berit, Børge, Ero, Leendert, Line and Tobias) which has provided both technical and social input along the way. The same goes for the rest of the NSEP-crowd who provided a lot of good discussions over lunch on topics ranging from fair-trade coffee to information security... Inspiring!

Last, but certainly not least; a final thank you is in order to Liv, who has supported me in her own way, regularly making me painfully aware of how unimportant location systems and other nerdy issues *really* are in the big picture.

Trondheim, 18. July 2009

Andreas Dypvik Landmark

Table of Contents

Abstract i

Preface iii

List of Figures xi

List of Tables xiii

List of Abbreviations xv

1. Introduction 1

  1.1. Background and Motivation 1

  1.2. Research Questions (RQ) 2

  1.3. Project Involvement 3

  1.4. Scope 4

  1.5. Structure of Thesis 5

2. Context Aware Computing 7

  2.1. Why Context-Aware Computing? 7

  2.2. Context – Becoming Aware 8

  2.3. Context-Aware Applications 9

  2.4. Location as Context 10

  2.5. (Computer) Architectures for Context-Awareness 11

  2.6. Summary 13

3. Location Systems 15
3.1. Terminology 15
3.1.1. ISO/IEC Terminology 16
3.1.2. Location as a Concept 17
3.2. Location System Characteristics 20
3.2.1. Localisation Techniques 21
3.2.2. Absolute or Relative Location 22
3.2.3. Quality of Information 23
3.2.4. Distributed System 25
3.3. Indoor positioning 26
3.3.1. Novel Techniques For Indoors 27
3.4. Location System Technologies 28
3.4.1. Beacons 29
3.4.2. Satellite Based Infrastructure 30
3.4.3. Utilising Existing Infrastructure 31
3.4.4. Special-purpose infrastructure for Location 33
3.4.5. Hybrid systems 35
3.4.6. Alternative approaches 35
3.5. Summary 36
4. Sonitor Indoor Positioning System 37
4.1. Description 37
4.1.1. Server Software 38
4.1.2. Base stations 39
4.1.3. Detectors 39
4.1.4. Tags 40
4.2. Installation, Tuning and Calibration 41

4.3. Sample deployments ... 43

5. Operational Capabilities ... 45

5.1. What are Operational Capabilities ... 45

5.2. Dimensions ... 46

5.2.1. Granularity and Amounts of Infrastructure ... 46

5.2.2. Resolution ... 48

5.2.3. Concurrency ... 49

5.3. Operational Capabilities ... 49

5.3.1. Method ... 50

5.3.2. Findings ... 51

5.4. Operational Capabilities as an Assessment Tool ... 53

5.5. The Need for Additional Dimensions ... 54

5.6. Summary ... 55

6. Location System Testing ... 57

6.1. Laboratory ... 57

6.2. Experiment Design ... 58

6.3. Baseline functional tests ... 59

6.3.1. Individual tag strength ... 59

6.3.2. Detector Consistency ... 63

6.3.3. Range and Angle of Arrival (AoA) ... 64

6.3.4. Tag Concurrency ... 67

6.4. More Functional Tests ... 72

6.4.1. Room-Level Precision and Accuracy ... 72

6.4.2. Zoned Room Precision and Accuracy ... 73

6.5. Evaluation of Test Results ... 78

6.5.1. Dubious Strength Measurements ... 78

6.5.2. Consistency 79

6.5.3. Range and Angle 80

6.5.4. Concurrency 81

6.5.5. Detector Handover 83

7. Evaluation 85

7.1. From Theory to Capabilities 85

7.1.1. State of The Art 86

7.2. From Experiments to Capabilities 86

7.2.1. Applying results to the Matrix 88

7.2.2. Granularity and Size of Infrastructure 88

7.2.3. Required Number of Tests 89

7.3. Value of Such an Assessment Tool 89

7.4. Cost 90

7.5. How to Utilise the Method in Practice 90

7.5.1. From Purpose to Location System 91

7.5.2. From Location System to Opportunities 92

7.6. Summary 92

8. Conclusion 93

9. Further Work 95

References: 97

A. Descriptive Statistics for Tag Consistency 101

30-Second P-Tags: 101

10-Second P-Tags: 101

7/10-Second E-Tags: 102

B. Descriptive Statistics for Range and Angulation 103

C. Purpose-made scripts 104

For reformatting to a more Excel friendly format: 104

For calculating Inter-Arrival Times(IAT) 105

List of Figures

Figure 1 A location Taxonomy (From Dobson[16]) 18
Figure 2 Sonitor Base Station (right) and a Detector (left) 39
Figure 3 Various Sonitor Tags. E-Tag (inset), two P-Tag cores and two P-Tag shells (open and closed) 40
Figure 4 Floor plan of the NSEP Usability Laboratory 57
Figure 5 Smoothed (2 Step Average) Time Series Plot of 10-second P-Tags (Left) and 10-second E-Tags (right) 61
Figure 6 Box plots of various tag strengths 62
Figure 7 Tag placement illustrated as seen from above 65
Figure 8 Line plots of Strength versus Distance (left), with a polynomial regression (right) .. 66
Figure 9 Radar-plot of strength versus angle and distance 67
Figure 10 Illustration showing the effects of CSMA on IAT 69
Figure 11 Expected number of detections versus measured detections for 2-minute 70
Figure 12 Expected number of detections versus measured detections for 5-minute 71
Figure 13 Pulse plot 71
Figure 14 8-tag (left) and 6-tag (right) Zone Accuracy Time Series Plot 75
Figure 15 Zone for Tag 60464 Over Time 76
Figure 16 Time Series Plot of Strength with Spike 78
Figure 17 Box plots of IAT over 2-minute windows at T=0min, T=8min and T=12min, group by Tag ID 83

List of Tables

Table 1 Scale of granularity and infrastructure.. 46
Table 2 Scale of Resolution 49
Table 3 Operational capabilities against granularity requirements.. 52
Table 4 Matrix of Granularity and Resolution from Table 3 53
Table 5 Room list 58
Table 6 Descriptive Statistics for Detectors 64
Table 7 Tally of detected locations for 8 tags spread over three rooms 73
Table 8 Statistics for non-calibrated zone.. 76
Table 9 Statistics for zone calibrated for equal average strength 77
Table 10 Results from experiments 87
Table 11 Matrix with Superimposed Test Results.. 88

List of Abbreviations

AoA	Angle of Arrival
API	Application Programming Interface
CLC	Centralised Location Computation
COSTT	Cooperation Support Through Transparency
CSMA	Carrier Sense Multiple Access
EM	Electro Magnetic
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSM	Global System for Mobile Communications
IAT	Inter-Arrival Time
IPS	Indoor Positioning System
IR	Infra-red
LLC	Localised Location Computation
RF	Radio Frequency
RFID	Radio Frequency Identification
RTLS	Real Time Locating System
TDoA	Time Difference of Arrival
UMTS	Universal Mobile Telecommunications System
WGS	World Geodetic System
WI-FI	IEEE 802.11 Wireless Network

1. Introduction

This introductory chapter provides an introduction to the problem as well as an outline of the background and motivation for taking on this problem, and the context out of which it grew out of. Further it includes a brief description of the scope of the thesis and finally an overview of the overall structure of the thesis.

1.1. Background and Motivation

The rapid decline in the price of location sensing equipment, coupled with location sensing capabilities being built into more and more equipment, have increased the demand for location systems and location based services. Hence the development and use of location systems are receiving increasingly more attention from a variety of industries. A recent report[1] predicts that the market for real time locating systems (RTLS) will increase from a global revenue of approximately $153 million in 2009 to $2.58 billion in 2019, with the number of suppliers rising from approximately 50 today, to over 200 by 2014.

Locating objects indoors has posed a challenge to the technology and methods used by the more mature outdoor positioning systems. Indoor systems are often tasked with tracking of rather large numbers of entities within small and confined spaces. This has created a slightly different set of requirements from the corresponding outdoor systems. This has led to the development of a set of systems specifically designed to tackle indoor positioning, often referred to as an indoor positioning system (IPS).

IPS technology is also seen as promising by the healthcare sector, as a potential way of increasing efficiency and quality as well as potential for new opportunities all together.

Examples of applications from the healthcare sector includes tracking vulnerable people[2], automating the monitoring of processes[3], reducing the risk of human errors in surgery[4] and avoiding wrong-patient-wrong-location issues[5], to tracking and keeping inventory of expensive equipment[6].

The applications of IPS in healthcare are still in its infancy. The emphasis has been on numerous unconnected pilots, proof-of-concepts and applications with narrow scopes. The adoption of IPS has very much been driven by one or a few specific needs rather than a clear and good overview of the whole picture. To successfully apply an IPS system that is realising return on the investment requires a thorough understanding of the technology. There are also examples in literature of well founded applications with a sound purpose which have failed to realise their full potential because of this mismatch between the required capability and the design of the installation[7]. For succeeding in implementing a location system it is important to understand how basic characteristics of the system (coupled with the effects of the environment it is installed into) affect the contextual information that can be extracted, as well as adapting the size of the installation to the correct level in terms of what information is required. It is important to *verify* the required capabilities, as well as *validate* whether a potential system can fulfil these requirements.

The purpose of this Master's Thesis is to investigate how to assess the required capabilities of a location system and to investigate how these capability requirements can be assessed experimentally without access to a full scale implementation.

1.2. Research Questions (RQ)

Given this problem definition there are two obvious research questions that formulate themselves, namely

> *RQ 1. What operational capabilities need to be understood before designing an indoor location system implementation?*

This is the verification step of the process, what is the potential variation between different systems and how is this important in the design of such an implementation. This is sought to be explored through use of example implementations from the domain taken from literature explored using context-awareness theory.

RQ 2. How can these operational capabilities be assessed by experiment?

This represents the validation step of the process where we seek to establish a set of tests that can be used to assess a location system for these operational capabilities through experiment. Using the Sonitor IPS equipment available, the devised tests will be evaluated against a real system.

RQ 3. How can one predict what impact the size of infrastructure will have on the system's operational capabilities?

The third research question has seemingly significant overlap between with the second question. However, the issue addressed by RQ3 is of such a nature that it is hard to fix experimentally prior to a full scale implementation. Therefore special attention is paid to how to be able to obtain a measure, or even predict how the size of the infrastructure is going to impact the operational capabilities of the system.

1.3. Project Involvement

The work behind this Master's Thesis is performed as a part of a larger coordinated research effort called ***C**o-**O**peration **S**upport **T**hrough **T**ransparency (COSTT)*. COSTT is a multi-disciplinary research project funded by the Norwegian Research Council. Some overall objectives for the COSTT project is[8]:

- "To enable flexible, 'Just-in-time' coordination of work in a highly collaborative and dynamic work environment"
- "To achieve this by creating a shared work space that gives all the actors involved in the collaboration real-time insight into the work process, e.g. its progress and possible deviations from the expected course."
- "To derive this insight automatically from samples of data obtained from the work environment by means of sensoring and monitoring devices"

This Master's Thesis falls in under the third objective of the project and is focused on acquisition of environment data by use of physical sensors – more specifically acquisition of location data by use of ultrasound based location sensors. The involvement in the COSTT project has also had an impact on the choice of theoretical approach. The overall idea of the COSTT project is very much geared towards a context-aware system, drawing upon theory from both context and context-awareness as well as computer supported cooperative work disciplines.

The COSTT project also has also established relations to several partners – ranging from academic research partners, health care institutions to industrial partners. Some of the materials used for this research have been provided by these partners, notably the Sonitor Technologies' *Indoor Positioning System*.

1.4. Scope

This thesis has by way of theory adapted a Context-Aware Computing approach to location systems. There are several other probably equally interesting concerns and aspects that could have warranted research and be assessed using different theoretical backgrounds. By implication this thesis will therefore not discuss non-technical issues such as privacy concerns and the ethico-legal issues involved with location systems that arises when deployed in the workplace.

Further this is not an electrical engineering approach to positioning, while new hardware and methods for locating objects is still very much an active field of research it is well beyond the scope of this thesis. This thesis will focus on the issues that are important when implementing a location system as the source of context information for a computer system. This means that technical challenges are explored from their implications for context gathering and what can be inferred from the location data obtained from the system.

Similarly there is a lot of interesting research going into areas such as Wireless Sensor Networks (WSN), sensor-web enablement (SWE), Sensor Markup Language (SensorML) and other sensor frameworks. All of these as well as the challenges with the current explosion in the number of available, and interconnected sensors, are also beyond the scope of this thesis.

1.5. Structure of Thesis

The first two chapters of the thesis contain the theoretical background for context-aware computing as well as more specifics on location systems. Chapter 2 examines some theory on context and context-awareness with a particular focus on location, and location as a contextual clue. Chapter 3 goes into more detail on the particulars of various location system technologies and common properties and techniques used in location systems.

Chapter 4 is devoted to a closer look at the Sonitor Indoor Positioning System which be used as the system for performing the testing in Chapter 6.

Drawing on the theoretical background from Chapters 2 and 3, Chapter 5 describe operational capabilities required from a location system, as well as literature describing the use of location systems in the hospital. The goal is to examine which operational capabilities needs to be understood when considering and later assessing a prospective location system.

In Chapter 6, using the location system described in Chapter 4, there is a description a design for experiments to assess the capabilities identified in Chapter 5. Based on this design the experiments are performed and the results are presented. Chapter 7 discussed this experimental approach to the assessment of capability. This is followed by overall concluding remarks from the Thesis work in Chapter 8 and the prospects for further work in Chapter 9.

2. Context Aware Computing

As set out in the project description, the overall goal of the COSTT project is to create a context-aware coordination software tool. This also frames the theoretical background for this thesis, which has adopted an approach to location data as a source of contextual information.

In this chapter the concepts of context and context-awareness will be explored with an aim of approaching the theoretical foundation for why location (which is particularly important to this thesis) is such an important element in context-aware systems. There is also a short discussion on why context-aware computing is a viable approach to the problem. To develop the context concept, relevant literature was researched and discussed with particular focus on location-aware systems.

2.1. Why Context-Aware Computing?

The continuous sensing of physical events - be it access card swipes or a more sophisticated location system placing people in a virtual model of a location, creates a well of location information about entities and their relationships. However, the information gathered can be used for various purposes and for a wide variety of applications. The methods used for collecting this information, including the limitations imposed by these methods can create important restrictions on how this data can be interpreted. Additionally it is important to be aware of what cannot reliably be inferred from this data.

In one extreme, location systems can be classified solely based on their physical properties. For instance, what is the best spatial resolution obtainable and under which conditions can these levels of operation be achieved. That would provide ample facilities for comparison between functionally equivalent systems. However, finding completely functional equivalent systems is difficult as the properties of the various technologies involved differ a great deal.

On the other extreme one can attempt to approach location from a more functional perspective, what is an appropriate use for it and what can it do regardless of the various properties of the system collecting it.

2.2. Context – Becoming Aware

Context is discussed in a wide variety of research disciplines, from linguistics to informatics. Concentrating on informatics, different disciplines adopt slightly different definitions and uses of context. In this Thesis we draw upon the definition, or rather definitions, proposed in the Human Computer Interaction discipline including the somewhat more specialised "context-aware computing" environments.

Context is an everyday concept and most people have a pretty clear notion of what it is and use it, yet most will find it difficult to specifically explain the concept of context. This is reflected in research, definitions of context range from definitions by example, definitions by synonyms, to more abstract definitions trying to elucidate and define the concept and its relations rather than specifying its instances.

Several researchers have attempted to reach a definition of *context*. Dey and Abowd in their article "Towards a better understanding of Context and Context-awareness" discussed in depth the fact that context is hard to conceptualise. Dey and Abowd put forward a definition of context as: *"Context is any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."*[9].

A decision can seldom be made on the basis of little or no contextual information. To enable a decision-maker either as a computerised algorithm or in human form to make a decision, there is some context that plays a part in the decision process. Using an everyday example such as trying to pick out a set of clothes when dressing up in the morning, we can see how the definition of Dey and Abowd is rather indistinct when using it to identify what the context of the clothing decision is. From their definition: *"Context is **any information** that can be used to characterize the situation of an entity. [..]"[9]* (emphasis added). This gives a rather coarse definition of what is context and what is outside the context of the decision. Winograd makes a clearer distinction between context and what he calls setting, rather than context; "Something is context because of the way it is used in interpretation, not due to its inherent properties"[10]. The example given is for instance whether the identity of the person using a computer is context or if the identity given to the computer by the user is the context. Extrapolating back to the original example of dressing up, Winograds more specific definition excludes anything not actually a part of the clothing decision-making process, but still in the vicinity of the person clothing himself.

Humans obtain context from a wide range of sources, ranging from prior experience to current and unfolding physical events in their presence. Having computers collect, organize and interpret the same amount and diversity of information is difficult to imagine. Humans are good at adapting to changing conditions and adopting novel practices based on previous experience. Computers on the other hand needs to have reasonably predetermined patterns to observe and act accordingly to, and does not really have the ability to reason in unknown territory or in face of unexpected deviations from the expected norm.

2.3. Context-Aware Applications

Context-awareness in applications is not something new. In its simplest form, users deal with context-awareness in computing every day. As Winograd points out in Architectures for Context, "What will happen when you hit the key marked 'backspace' on the keyboard?"[10], the answer is dependent on which application is currently active, i.e. the action is dependent on the context of the key press. With graphical user interfaces there are large amounts of context included in human computer interaction, which most users intuitively adapt to and exploit to go about their daily business. To say so, the function of the backspace-button has become second nature and proficient computer users do not actively or explicitly consider its context before pressing it. Dey and Abowd describes context-aware applications as "A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task"[9].

In interaction between computers and humans (or really between two interlocutors of any kind) a successful interaction (or conversation) is dependent on the parties being able to interpret each other's intentions. A shared understanding of intention is dependent on mutually available context.[10]

2.4. Location as Context

Location is undoubtedly an important part when attempting to "characterise situation of an entity". In the review by Dey and Abowd[9] they identify several definitions-by-example of context, four individual pieces of work which all lists location as one of few elements (others include identity of surrounding people, time, orientation, objects and people) [11-14].

Identity, (current) activity and time are all contextual information readily available to a computer. The identity of the user (or even recipient) is usually known through manual input (as simple as a username/password-combination to access it, or more elaborate schemes utilising biometrical information such as fingerprints, etc.), likewise the activity that an event occurs in is often known through what program is currently active (or in focus) on the computer. The progression of time, as well as current time, is something computers usually can access directly through built-in circuitry. Location on the other hand is something that cannot be captured directly by an ordinary computer. For a computer to be able to sample something from the physical realm it has to have some input from some kind of sensor, generally speaking, a location system.

The implications of location as a contextual clue, is larger than the obvious "knowing where objects are". It can often help characterise a situation. For instance if the surgeon is not present in the operating theatre, a knowledgeable surgical coordinator can infer that the surgery has not started. Depending on the access to other contextual clues, such as the location of the patient, anaesthesiologist, or even the surgical plan for the department, etc. a coordinator could potentially infer even richer details, i.e. not only the fact that the surgery has not started, but if it has been postponed or even cancelled, etc.

The "surgeon-example" above is dependent on the coordinator being able to access, process and conclude based on some input of contextual clues. Similarly one could imagine that a computer given ample amounts of contextual information of varying accuracy could be able to infer certain characteristics about its surroundings, similar to those that the surgical coordinator does.

2.5. (Computer) Architectures for Context-Awareness

The problem, and possibly a part of the reason why advanced context-aware computing has not become main stream is the lack of coherency and interoperability between various implementations and sensor equipment. Each application that wants to sample location information from a sensor array would have to implement support for the sensor hardware directly. This leads to a situation where one has to re-implement or add additional program code for each system that is sought to be supported by the software. This often referred to as the “silo-model” of software engineering. Each application creates its own silo of supported functionality potentially spanning from the hardware level up to the user interface with little or no re-use or exchange of data or functionality. Thus each application appears as an allegorical self-contained silo.

Context-awareness is gaining ground and becoming increasingly more popular with mobile phones becoming full-fledged computer systems as well as small laptop computers becoming more and more popular. This has led to a breakdown of the traditional silo-model of sensor integration where sensors are integrated for one application with one purpose. In modern computing there is trend towards generalized (and open) Application Programming Interfaces (APIs) to enable reuse and interoperability[1] - context-awareness is also moving in this direction. One example of this trend is Microsoft’s introduction of the Windows Sensor and Location Platform in their upcoming release of the Windows operating system (dubbed “Windows 7”).

The idea behind the Windows Sensors and Location Platform is to give a unified API for accessing sensor input. Their current technology showcase includes a sensor-board including

---

[1] As reflected in other trends such as “Web 2.0” and “Mash-ups” – technology applications at the intersection of service oriented architecture (SOA) and open, in terms of specification, web-accessible APIs.

an ambient light sensor, accelerometer and “human presence”/touch-sensors for detecting human touch. With the correct driver support the idea is that the ambient light sensor can be tapped for information on lighting conditions which again can influence the amount of backlight and contrast used when displaying information – to automatically improve the reading experience when moving the computer between different environments.

While this ambient lighting functionality is rather simple feature, it is a nice example of context-aware computing as well as something that probably would not have been realised without a common API. Implementing such a minor feature would have been a large undertaking if the sensor-support as well as the display driver support had to be implemented as well. Given the available API it is more a case of implementing the *glue* that brings the two APIs together. This makes for easier implementation for both application developers (by abstracting away hardware thus making it easier to develop generic applications) as well as users (not having to configure a device in every single application he or she wishes to use it in).

The Location Platform is intended to be sensor agnostic from a developer point of view, making it possible to tap location information without knowledge of whether the source is GPS, WI-FI triangulation or even user-entered data. The Location API will abstract and “choose” the better source depending on location (for instance WI-FI for indoor, GPS for cross-country trips outdoors). There are also some rudimentary[2] controls for addressing privacy concerns with regards to access to sensor input.

---

[2] The ability to control source access depending on user as well as differentiating between sensors for different users – allowing some to have sensor input from accurate sensors whilst others only access to more coarse-grained sensors.

http://www.microsoft.com/whdc/device/sensors/default.mspx , Accessed 2009-05-11

2.6. Summary

As ubiquitous computing continues to grow in popularity and devices become smaller and increasingly more mobile, the context-enabling of software will become more and more important. The use of context-aware devices will both support users in performing current tasks more efficiently, increase the usability of the current tools, as well as opening up possibilities that just is not possible with the current set of systems.

Becoming aware is very much dependent on sources informing the system of meaningful changes in context, as they happen. Currently there is a lack of wide-spread generic APIs for abstracting context-information. In the next chapter there is more discussion on why such APIs are difficult to realise and how the diversity of context, and in the scope of this thesis, location are rich concepts and therefore difficult to create good generic APIs for.

Location is a source for a lot of different types of contextual information that can be utilised for different purposes and it is therefore important to be able to capture the location of both people and things that falls within the context of a system to enable it to draw upon this information.

3. Location Systems

As briefly discussed in Chapter 2, location is a central element in context and context-awareness. Location is used as an everyday concept by people everywhere and a closer examination of this concept is warranted before attempting to capture it and representing it digitally.

The aim of this chapter is to explore the main concepts and notions important for location systems and provide a background to later discussions about various location systems.

This chapter begins with a discussion on the abstract concept of location. This is explored through consulting relevant literature from both pragmatic as well as theoretical approaches to location and location systems. This investigation also includes a discussion of key terminology and methods related to location and location sensing. These are clarified in the beginning of this chapter. The latter part of the chapter is devoted to a description of the technology alternatives and an exploration of the state of the art in location system technologies.

3.1. Terminology

In everyday use *location* and *position* is almost synonymous. One can be both on location and in position which is semantically the same thing. As with many other terms, these two words have a more specific meaning as representatives of two concepts in the positioning and sensor world.

Location is, as such, a concept with a richer or larger set of possible values. “The office” could be a valid location, but it can hardly be described as a position without some frame of reference further describing who’s office, where this office is, possibly related to some temporal aspect (one person can have several offices).

This thesis has adopted, based on taxonomies by both Hightower and Mannings, the term *location* and hence the term *location sensing* for the act of deriving a location based on some input. L*ocation system* refers to a system describing and performing such operations (sometimes including some transformation and interpretation of the sensed data). Position, on the other hand, is a physical point (or place) which can be described as a location or one can refer directly to it using some general form of coordinates or frame of reference.

These terms could to some extent be used interchangeably with terms such as *Indoor Positioning System (IPS)*, for instance as used by Sonitor. Ironically the system referred to as Sonitor IPS is in reality very much an Indoor *Location* System and as such does not provide position data, but rather location data.

3.1.1. ISO/IEC Terminology

ISO/IEC standard 19762 defines several terms for Real Time Locating System (RTLS)[15]. It is worth to note that RTLS as defined by ISO/IEC is a stricter definition of location system than the one used in this Thesis and in a lot of literature in general. Notably ISO/IEC explicitly excludes certain types of passive systems, cell proximity systems as well as systems relying on beacons transmitting without any active interaction with the tags. These systems are included in the looser definition of location system, as applied in this Thesis.

For further clarification the following terms, as defined by ISO/IEC are used in the Thesis, with the exception of a different word for the same concept for the terms *Reader* and *Transmitter*:

Server – computing device that aggregates data from the readers and determines location of transmitters.

Reader – device that receives signals from an RTLS transmitter.

In this Thesis the ISO/IEC Reader will be referred to as *Detector*. In other literature it is also commonly referred to as a *Receiver*.

Transmitter – (RTLS) active radio devices that utilize the specified RTLS protocols.

In this Thesis the ISO/IEC Transmitter will be referred to as *Tag*.

Infrastructure – (RTLS) system components existing between the air interface protocol and the RTLS server application programming interface (API).

3.1.2. Location as a Concept

For any location system, a central concept is, of course, *location*. A location system is supposed, by definition, to derive some *location* based on some input. To enable a comparison of systems a common understanding of the concept of location is required. As we will come to see, location is a surprisingly rich concept with several diverse facets.

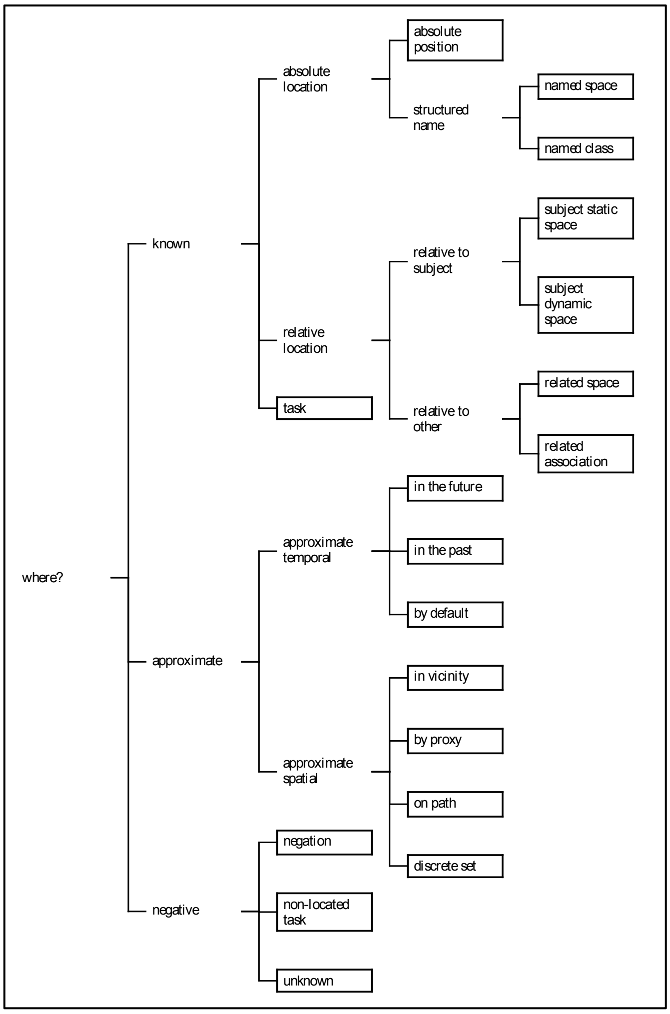


Figure 1 A location Taxonomy (From Dobson[16])

Dobson in his paper “Leveraging the subtleties of location”[16] deals with precisely this very question, what is a location. Attempting to uncover some of the richness that lies in the concept of a location he puts forward taxonomy for location, wherein he identifies 18 different answers to a location request. His taxonomy is illustrated in Figure 1. Dobson divides location into three separate branches; known location, approximate location, and negative location.

Known location is the simplest answer to a location query. Dobson covers the possibility for answers in terms of absolute locations, relative locations as well as answering in terms of tasks. That is, replying with the task the sought person is performing rather than the location he is in, thus leaving it up to the recipient to infer more information from the task.

As shown in the figure, approximate locations can refer both to the temporal and the spatial domain, i.e. a location can also be given as “Robert was at home at 10:00” which is in effect contextual layering adding a temporal layer to the location – the precise location of Robert is in fact not known, but factoring in contextual clues such as the location he was at a given time it can be combined with other pieces of context (is he travelling by car?, where is he headed?, etc) to derive a relative or even absolute location (though perhaps with less than absolute accuracy).

An interesting inclusion is that of *negative location*, i.e. where the located object is not. While not obviously containing directly interesting information, a lot of location-dependent systems can be viewed as equally dependent on non-location. For several purposes it is just as informative to know that the person is “not at home” or is “out of office” rather than knowing where he currently is vacationing or if he is just visiting the dentist. From a privacy perspective negative locations can also solve certain privacy issues by not disclosing specific information, but rather removing the actual location that is interesting from the domain of discourse while at the same time leaving the domain so large that it is uninteresting for other purposes.

There are several implications from having such a rich taxonomy for location. As Dobson briefly touch upon in the latter part of the paper, having such a rich and general taxonomy means that a system striving for generality would be infeasible given the current state of the art in software engineering. The span of possible answers crosses several types of context (temporal, spatial, etc.), just the collection of them would be problematic as well as the fact that reasoning over such a diverse model would require a rather extensive engine. However, as Dobson points out most current systems have rather narrow or specific use cases and tends not to deal with a full general model of location.

Other writers have approached the same problem and devised simpler taxonomies, such as Indulska and Sutton. In their *Location Management*-framework[17] they identified three types of location information; Physical, and Virtual or Logical. This separation is based on what type of sensor or data is used for generating the information. Physical location is data based upon information sensed from the physical realm, i.e. by way of sensors or actuators, etc. Virtual location, on the other hand, is information based on such as mouse or keyboard activity or an electronic calendar, or even content of e-mails, etc. They also introduce the concept of logical location; that is location information logically inferred from other sources.

For instance the recorded fact that you unlocked a door with your ID-card places your location in the room.

Indulska and Sutton do not make a clear distinction between what is virtual location information and what is logically inferred location information. Others have made more clear distinction on the amount of interpretation and synthesis performed to reach a location. For instance calendar information might contain a described location directly and involves little logical inference apart from time, while synthesising utilisation of telephone and access card use for instance is closer to a logical inference to derive a location.

There are also similarities between the two taxonomies by Indulska[17] and Dobson. However Indulska created a taxonomy based upon collection method while Dobson's taxonomy is ignorant to how the information came about. Dobson does not directly discuss the collection of information, but emphasises that it is the synthesis of sensed location with information from *"the widest possible set of sources"* [16] that is important.

3.2. Location System Characteristics

It is difficult to characterise a generic location system in detail. Various systems attempt to derive locations using a wide range of techniques, technologies and sensor equipment, each with its own advantages and limitations. This also means that for a system designer attempting to solve a particular problem there is an extensive set of questions that needs to be answered to enable a decision on what features – or which system – to pursue.

Hightower and Borriello in their "Survey and Taxonomy of Location Systems for Ubiquitous Computing"[18] identify eight important characteristics or properties that apply to most location systems (due to the span of location systems, not all properties apply equally to all systems).

- Physical position and symbolic location
- Absolute versus relative system of locations
- Localized location computation
- Accuracy and precision
- Scale

- Recognition
- Cost
- Limitations

Their characteristics are aimed at covering physical sensor based systems. However, as previously mentioned, location sensing is not limited to sensing the physical realm. It is possible to “sense” location based on virtual sensors, for instance looking at network traffic on a work station to see if there is someone present. It is also possible to infer locations logically, i.e. “sensing” a location using logic and drawing information from a fact base such as calendars or even emails. In the scope of this Thesis it is the physical realm that is applicable and the remainder of this section will be devoted to presenting different issues that is central to location systems on the whole.

3.2.1. Localisation Techniques

There are, in principle, three main techniques shared by most systems for determining location based on some input; multiple input combined with a localisation algorithm, proximity and scene analysis[19]. Of these, the first two are perhaps the more commonly found in practise.

The main algorithmic method technique is usually attempting to construct triangles between the sought location and two known points of reference, using geometry to solve for the missing information. If the system sensors are able to measure distance to known points (lateration), the constructed triangle(s) are solved for known side lengths (trilateration, or multilateration). The other alternative (angulation) measures the angle of incidence between known point(s) and the sought location (triangulation).

Lateration can primarily be measured using *time-of-flight* (measuring the time it takes to transmit or move something that travels at a predetermined velocity between a known and the sought point), *attenuation* (measuring signal strength and correlating it to a predetermined curve of signal decline) or the theoretically simplest method *direct*, that is directly measuring the distance.

Angulation in two dimensions requires two measurements of angle and one known location. This is usually done either through having measurements of angle (relative to a reference vector, often perpendicular to a virtual line between the two stations) from two known stations with a known distance between them, or conversely using two antennae for angulation with a known distance between the two (local) antennae.

Trilateration and triangulation is sometimes used somewhat interchangeably in literature, most likely because the algorithms use the same principles, but the data required and the sensors to detect such data usually vary.

Proximity is basically detecting whether or not an object is near (for some appropriate value of near) a known location. This can be detecting physical contact between the object and the known location, detecting that a device is within range of some wireless network or observing automatic identification of objects which are identified by a device with a known location (and thus inferring proximity).

Scene analysis, the last of the three principal techniques, relies on identifying features in an observed scene or “landscape”. By identifying features of an object and comparing it to features with a known location (e.g. a “horizon”) one can map an object to a location. This is called *static* scene analysis, conversely in *differential* scene analysis one compares the change between different scenes using the difference to compute the movement of an observer. Scene analysis is not limited to visual imagery (i.e. video feeds, or pictures). It can also utilise other representations of landscapes, such as the presence of radio-transmissions which then forms a radio wave “landscape” which can be analysed and recognised.

3.2.2. Absolute or Relative Location

There are two ways of describing a known positive location – absolute or relative. An absolute description of location places a located object into a pre-established frame of reference. This means that any other sensor operating with the same set of references, which tries to sense the same location, will return the same location (this is of course ignoring differences in information quality such as differing resolutions, accuracy and/or precision). Examples of common reference systems are for instance the World Geodetic System (WGS) (which is commonly used for ordinary maps, map-making and navigation and perhaps the most well-known system for the ordinary user).

A relative location, on the other hand, gives the location of an object in relation to another object (or space). This could be seen as establishing a transient local frame of reference, which is linked to knowledge about a local object or space, whose location might also change over time. In comparison WGS, uses the Earth's centre of mass as the origin of its coordinate system, requiring a detailed and specific model of the earth to obtain accuracy. On the other hand, any significant change in the circumference or shape of the earth would probably have more wide reaching effects than having to update a few maps. While this enormous reference system is required if one wishes to navigate or measure details on a global scale, it is often unnecessarily complex both in use and algorithmically when only trying to locate an object in a house with a well-known layout.

For these, more local applications, relative locations can be appropriate. This could entail giving a location of object A as *"next to object B", "within space B"* or more detailed *"10-centimetres left of object B"*. It is obvious that this location description is only useful if all the users have a common reference. Without a common reference, one is left with questions such as *what is* and *where is* object B, the orientation of the system (or what the frame of reference is).

3.2.3. Quality of Information

There are several parameters pertaining to the (perceived) quality of the information provided by a location system. This thesis adopts the term *quality of information*, as used by Mannings, to collectively refer to the four basic properties: accuracy, precision, granularity, and resolution. While they represent different challenges and have different underlying causes, the effects of changes in these dimensions can impact a system in similar ways. In this section we will attempt an explanation of how they apply to location data.

The first two, *accuracy* and *precision* are in fact statistical measures of data quality. Accuracy is a measure of how close a sensed location is to the true location of the sensed object (i.e. the distance between the sensed location and the physical position). This is usually measured in units of length (i.e. metres). Precision on the other hand is a measure of how often one would expect to achieve a certain accuracy. Usually given in percent, for example a system can be said to be accurate to <10-metres at 95% of the time. Conversely one would expect the system to be inaccurate to this <10-metre measure on the order of 5% of the time.

As we can see accuracy and precision are interrelated, and there is often a trade-off involved meaning that improvement of accuracy can often reduce the precision and vice versa. One usually also operates with one measure of accuracy and precision for sensor-level and a different one for system-level. One reason for this is that some systems make up what they “lack” at sensor-level by clever use of algorithms. S*ensor fusion*, that is using a combined result from several sensors, can provide a sensed location which is of better accuracy and/or precision than a single sensor is able to provide.

Optimally a system with high level of accuracy and precision is beneficial, but acceptable levels of accuracy and precision depend on the application of the system. For different systems, different levels of accuracy and precision will be appropriate and acceptable. Trying to locate a migrating animal would probably not require much of accuracy nor of precision. A location on the order of kilometres would possibly be enough. On the other hand, a nurse trying to locate a patient or a piece of equipment, knowing a location on the order of a kilometre is in most cases probably useless. There is also often a correlation between cost of system infrastructure and accuracy/precision.

Granularity is used to refer to the smallest space that a location system can distinguish from another space – i.e. with room-level granularity a system is able to separate one room (i.e. the space bounded by a room) from another room. It is not able to discern whether the located object is close to a wall or if it is in the middle of the room. This is also sometimes referred to as resolution in literature, but it is important to make a distinction between a measure of spatial detail or granularity, and a measure of temporal detail or resolution.

Resolution in this context is a measure of the update frequency of the system, or how often the system can locate an object – i.e. it is a temporal measure of detail. For instance a system based on beacons transmitting once per hour can be said to have a temporal resolution of one hour. This means that this system is not guaranteed to detect a location where an object resides for less than one hour (depending on other parameters of information quality, residing in the same location for over an hour might guarantee detection). It is common to talk about “real-time” location systems; real-time here is relative to the application and use of the system and does not refer to real-time in the classical computer science/engineering sense. For a system tracking orderlies rapidly moving through large number of rooms, resolution on the order of seconds (or even lower, depending on the granularity of the location) would be “real time”. On the other hand, a system tracking primarily stationary equipment does not require the same resolution to be of equal value as a tool for tracking the movement in “real-time”.

3.2.4. Distributed System

Quite a few location systems are in fact distributed systems, in the computer science sense. They often contain several processing units requiring communication and cooperation to reach an answer. This means that several location systems are faced with classical challenges faced by distributed systems such as synchronisation issues where the system relies on a global notion of time (commonly seen in system relying on time-of-flight measurements).

There are also issues with where does the system compute the location, centrally or localised (meaning at the location, on the tag, etc.). Central Location Computation (CLC) gives a centralised control, potentially cheaper (in terms of lower computational power required) tags or user equipment at the expense of a less scalable architecture. The central computational facility has to scale with the number of tags and/or users. Localised Location Computation (LLC) on the other hand lets the tags calculate their own location based on input from infrastructure. This reduces the computational requirements on the central part, but also reduces control over the users. For instance GPS is a globally LLC based system, that means that the controller of the infrastructure (in this case the United States Government) cannot see who uses the system or their location. This also means that GPS can scale almost to infinity without increasing the load on the central infrastructure (in this case, the satellites and ground stations).

In terms of privacy and security concerns, LLC as a principle can potentially return a lot of the control over information to the person carrying the tag or equipment rather than those controlling the central infrastructure.

3.3. Indoor positioning

While the challenges are similar, there are important differences or variations between outdoor and indoor positioning. Outdoors are usually unbounded by physical boundaries. Some features of landscape (land-sea, rapid changes in altitude) might make it difficult to traverse from one space to another using one particular form of transport, but there are few boundaries that are impossible to cross using any form of transport. However, indoors there are physical boundaries that a system needs to be aware of and take into account. Indoor movement is for instance constricted by walls, floors and the doors linking these spaces.

Indoors there is a higher emphasis of identifying the correct space in which the tracked object resides in, i.e. which room is it in rather than the coordinate. It is therefore important to make sure that the virtual boundaries as represented in the location model of a system coincides with the physical boundaries that constricts the real world.

For both indoor and outdoor systems range and line-of-sight is usually a challenge. The response to the challenge is usually met with different answers. Where outdoors it is often countered by choosing frequency bands where wave propagation is less hindered by the current obstacles (for instance choosing low frequency solutions such as LORAN-C over GPS can make a location system more resilient against line-of-sight issues and so called "space weather"-issues, and vice versa. These outdoor systems usually operate over large distances and are mainly designed to permeate obstacles rather than to accommodate them. Indoors, this permeation of obstacles would mean that the physical boundaries would not be respected and as such a different approach to these challenges has to be deployed.

When applying this type of technology to indoor scenarios two issues crop up.

1. Indoor positioning is rarely line-of-sight positioning. Obstacles and unintended reflections are the rule rather than the exception. Physical boundaries are key elements in an indoor location model and are often fundamental building blocks when attempting to describe location in terms of a local frame of reference. Using obstacles as reference also coincides nicely with the physical perception of the users.

2. A global/common frame of reference is rarely a goal. A localised reference (or “floor-map”) is often more valuable than global coordinates or a standardised description of location.

These short comings coupled with the fact that the current large scale outdoor systems rarely function appropriately indoors without augmentation means that there is a niche for purpose-built and locally deployed location systems designed exclusively for indoor challenges.

3.3.1. Novel Techniques For Indoors

In light of the general problems faced by large scale systems a different set of technology rarely used outdoors has grown to meet the needs for indoor positioning. In meeting this challenge the main contenders in terms of medium are optical, ultrasonic, and to some extent electro-magnetic/RF solutions. As we can see microwave (as used by satellites, e.g. GPS) is missing from this list as it is not considered a viable option for indoor use (both health and technical issues arise).

When considering how these technologies deal with the alignment of physical and virtual boundaries these three solutions divide into two groups, with optical and ultra-sonic solutions in one group which cannot easily permeate walls or other reasonably solid obstacles while electro-magnetic waves being the other category that does permeate walls. While optical and ultra-sonic solutions are effectively stopped by walls electro-magnetic waves can traverse walls (depending on frequency and materials of the wall) with relative ease[3]. This means that less care has to be taken to constrict the detection to room-level view. The main difference between the ultrasound and optical systems, in terms of functionality, is in the robustness in signalling and the demands for line-of-sight.

Another issue, that mainly comes up in specific environments is also the interference caused by the signals broadcast by the location system. For instance in hospital interference between radio transmitters and medical equipment is taken very seriously. Consequences of

---

[3] As seen for instance in Wi-Fi deployments where one base station can cover several rooms.

unintended interference could be life threatening or occlude and distort results from diagnostic equipment. However, even with this widespread acknowledgement of the dangers of interference there is little published evidence supporting these claims. In fact most hospitals now have relaxed their policies on the use of radio transmitters in their buildings. There is however still some evidence that point to certain types of equipment still showing adverse effects from interference at close range [20, 21].

While this issue is widely used by manufacturers of systems relying on other transmission forms, there is an element of this with any technology. Both ultrasound and infra-red is used in medical imaging and diagnostics (though IR on a much smaller scale than ultrasound). This is also a two-fold problem; the main issue is of course the location system interfering with the day-to-day business of the environment it is deployed into (i.e. disrupting medical equipment), but conversely there is a problem the other way around with the environment disrupting the location system. Electronic equipment can and does emit "noise[4]", for instance fluorescent light fixtures can often create noise in ultrasound bands, the same goes for LCD/flat-screen displays.

In terms of published results, rudimentary searches in PubMed on the effects of ultrasound or infra-red interference on imaging equipment yielded no results. This is possibly due to the fact that the use of ultrasound and infra-red for other purposes is rather new and there has not been much research on it yet. It is however unlikely that new technology is completely without drawbacks, and as with everything else there is an environment-dependent risk/benefit analysis required to determine which system is more appropriate.

3.4. Location System Technologies

Technology for obtaining location or position data has been used almost throughout recorded history. From the apocryphal tale of the Three Wise Men of The Orient, navigating to Jerusalem aided by the star of Bethlehem to more modern and technological examples

---

[4] i.e. unintended transmissions not used for any purpose, in that respect adding to or even creating background "noise".

such as avalanche transceivers being used to locate people or equipment buried in snow. With the advent of computers the manual work involved in calculating position has been computerised and automated.

The methods used by computers to locate people and objects span a huge range of methods and use of different infrastructure. Imposing order on this large set of diverse technology is difficult, but a summary of the more common solutions and their applied technologies is included below. The summary does not try to impose a definitive classification of their technological merits, but more a loose grouping based on notable properties.

3.4.1. Beacons

Just as the apocryphal example above of the Three Wise Men navigated to Jerusalem by following the star of Bethlehem as their guiding beacon, sailors navigating close to land are also familiar with navigating aided by beacons, or more specifically lighthouses. Lighthouses are classic examples of beacons providing some information about location that a recipient (i.e. the captain of a ship) can interpret to make some prediction about the position of the ship. This position is usually relative[5] to some hazard such as underwater reefs, shoals or even placing the ship in virtual traffic lanes for vessels.

The modern versions of beacons are usually based on radio waves rather than visible light. Common for beacons is that they usually transmit information about themselves or their surroundings which a recipient can interpret to deduce their own location relative to the beacon. Even though they only transmit relative locations, a relative location can enable the calculation of an absolute location.

Beacon based technology is still very much in use, *LORAN-C* being a good example. LORAN-C is a short-wave based radio beacon system which still is used by ships worldwide for navigational aid. Aircraft navigation is also still based on radio beacons to help them “home

---

[5] Though using a sextant or similar equipment one can use angulation to calculate an absolute location based on the light from two lighthouses if the absolute location of the lighthouses are know.

in" on (the physical) landing strips or aligning their flight paths to (virtual) air corridors when traversing busy airspaces.

Most of the technologies described above are used for navigating oneself; however there are also examples of beacon-based technology applied inversely where the recipient has a "known" location and the beacon is to be located. An example of this is for instance the use in emergency equipment, allowing rescue services to locate persons in need by locating a beacon worn by the casualty. In difficult terrains this can be highly effective. One study of the use of avalanche transceivers in Austria found a significant reduction in median burial time from just over 100 minutes down to 20 minutes[22] by the use of avalanche transceivers. Similarly the "black boxes" of airplanes are usually fitted with underwater locator beacons that active when in contact with water and then transmits an ultrasound pulse to allow them to be found and recovered.

Beacons usually rely on triangulation or trilateration as a method for calculating location. In some cases just a single beacon, such as the black boxes or the avalanche transceivers. This does not provide enough input for direct calculation of location. The location has to be derived by moving the sensor equipment to form several points which then can be triangulated/trilaterated, or one simply moves the sensor equipment in the favourable direction (e.g. in the direction that increases signal strength) until one reaches the source of the signal.

3.4.2. Satellite Based Infrastructure

Satellite based location systems, more commonly referred to as Global Navigation Satellite Systems (GNSS), can also be seen as a pervasive set of beacons. These systems usually rely on the timing of microwave transmissions from satellites orbiting the earth to derive a location. These systems allow users with the appropriate equipment to determine their location based on these signals.

The most commonly used, and known, satellite-based location system in operation is Global Positioning System (GPS). GPS is perhaps the system most users think of when hearing the term location system. It provides near world-wide coverage at a minimal cost to the end user. GPS relies on having *time-of-flight* measurements from a minimum of three satellites (four for calculating altitude).

Other (all unfinished) examples of satellite-based global systems are Galileo[6] (funded and run by the European Union), GLONASS[7] (Russian undertaking) and COMPASS[8] (Chinese). COMPASS is a planned global extension of a regional satellite based location system.

Typical examples of GNSS-applications are the obvious navigational aids for both man and vehicles, tracking equipment or personnel. The latter is often called *Intelligent Transport System (ITS)*. Other applications include locative applications such as Brightkite[23].

Whilst the end-user cost for GNSS is usually restricted to the receiving equipment and rather negligible compared to the coverage area and functionality offered, the cost of the complete infrastructure is enormous. This is also the reason why system owners of GNSS systems are governments or even consortiums of governments. The deployment of such a system is not something easily undertaken by anyone. For example does a recent report[24] issued by the U.S. Government Accountability Office (GAO) indicate that the cost of keeping the GPS system operational is increasing to the extent that it is not sure whether the US Air Force (which maintains parts of the space infrastructure) will be able to keep maintaining and upgrading the infrastructure fast enough to avoid any service disruption.

3.4.3. Utilising Existing Infrastructure

Similar to beacon-based technology, there is also a thriving category of technology based upon using existing infrastructure (infrastructure primarily indented for other purposes) for location purposes. The obvious advantage for these systems is that the cost of deployment is often much lower than for a system requiring deployment of both infrastructure as well as tags. Since the cost of infrastructure often shows a correlation to the size of the space covered, being able to exploit large existing infrastructures that already covers large areas can prove cost-effective. Typical infrastructures used are cell phone networks (e.g. various GSM, UMTS, etc.) or television broadcasters, which both commonly have near complete coverage for populated areas.

---

[6] http://www.esa.int/esaNA/GGGMX650NDC_galileo_0.html Visited 01.03.2009

[7] http://www.glonass-ianc.rsa.ru/pls/htmldb/f?p=202:1:9467314884255617370 Visited 01.03.2009

[8] http://www.navchina.com/english/ Visited 01.03.2009

However, infrastructure built for supporting other purposes than location detection often have other requirements in terms of coverage and base/cell density that can make it difficult to achieve good or even consistent results.

Using existing infrastructure is not limited to wireless infrastructure. Wired telephone networks have been used for location information almost since its inception. Emergency services have relied on using proximity for locating calls when the caller cannot explain his or her location for various reasons. This was based on the location of the lines being known and a simple lookup of line-number in an address register. With the advent of mobile telephone technology in process of replacing land-based telephony this has led to an increased demand for location support in the existing mobile telephone network. As the density of various networks increases, the information quality of the location services usually increases too.

The methods applied for deriving location, varies with the characteristics of the infrastructure it attempts to exploit. For systems based on infrastructure with low density of transmitters a proximity-based location is often the only viable result. This places an object within the coverage of a transmitter, whose location is previously fixed giving a relative location (*"in the vicinity of …"*). For networks with higher density of transmitters, such as mobile/cell-phone networks, it is often possible to be in range of several transmitters which allows the application of triangulation/trilateration to derive at a more precise location. Again the system is dependent on having known locations for the transmitters, but it can give a much more fine-grained result than merely proximity to a (sometimes) large coverage area.

An example of technology that utilise existing infrastructure are, as already mentioned, positioning based on mobile phone networks as utilised by emergency operators (and often used in judicial matters for proofs). Other networks utilised is digital terrestrial television signals, for instance as used by Rosum[25]. Rosum has several methods for calculating location, for instance using digital television broadcasts, which already include time information which can be extracted and compared, as well as using additional information about the "virtual RF landscape" created by these transmitters. This is supported by regional servers that contains pre-established models used to inform and help the system calculate its location.

These technologies will most likely also become more important in the future as location based services become more widespread. The use of GNSS in dense urban areas is problematic due to the lack of sight created by the “urban valley”[9] where tall structures creates obstacles for a clear line of sight to the required 3+ satellites.

3.4.4. Special-purpose infrastructure for Location

Active sensor technology that requires infrastructure and sensors, transponders or tags attached to the objects to be tracked. This category is mostly dominated by systems intended for indoor use.

Traditional GNSS-systems usually perform poorly[10] (or not at all) indoors because of the lack of line of sight to their infrastructural beacons. Additionally, indoors a symbolic location of high accuracy is perhaps more valuable than a lower-accuracy absolute geo-referenced position, as usually produced by a GNSS. Similar arguments apply to several of the existing infrastructural-based systems. The accuracy required indoor is tightly linked to physical obstacles, such as walls. For this reason, there is a whole set of technology that has grown out of trying to locate objects indoors.

Naturally, indoor location systems operate over a much shorter range than their outdoor counterparts, each system is normally limited to one building or a campus of buildings. This allows for technologies that do not scale both in terms of clients, technology and infrastructure to outdoor applications. One prime example of one such application of

---

[9] “Urban Valley” is an expression used to describe the reception often found in cities and urban areas where man-made structures rather than terrain causes obstacles for obtaining line-of-sight, similar to what is experienced in the bottom of valleys.

[10] There are several technologies aimed at mitigating these difficulties – solutions ranging from the deployment of pseudo-satellites (pseudolites) that emulate or even replace the signals received from actual satellites hence creating its own set of beacons similar to other technologies whilst preserving the user terminals. This and other methods are often referred to as Assisted-GNSS (A-GNSS) including both pseudolites as well as transceivers and other methods for improving signal strength indoors. However, most of these also introduce their own set of problems (often related to problems with global clocks).

technology for indoor positioning which is impossible in outdoor systems is the use of Radio-Frequency Identification (RFID).

RFID is a tiny integrated circuit with an antennae-array which can modulate a radio signal through a principle called *modulated backscatter*. The technology has been around for the better part of the 20th century. Initially being used for identifying friendly military units on radar equipment and later used for electronic payment on public transport[26], embedded in passports[27] and used as a replacement for barcodes on merchandise[26]. Deploying a number of RFID readers, one can implement a system which can, through proximity (or even more advanced algorithms, see for instance SpotOn[28]) track an object moving through the covered building volume.

RFID can be either active (with an external power source, usually batteries) or passive (powered by the current induced by the field created when attempting to read a tag). Passive tags offer the advantage of having a near infinite life-span at the expense of being limited in terms of functionality.

Other systems (e.g. Olivetti's Active Badge[29]) have been based upon deploying matrices of infra-red (IR) sensors and equipping the objects (usually people) with IR-transmitters. Using IR-transceivers these tags can also be used for simple (low-bandwidth) two-way communication. The obvious drawback being the poor performance of IR under some lighting conditions as well as a strict requirement for line-of-sight (i.e. the tags cannot be occluded by clothing, etc). IR is light-waves and thusly cannot permeate walls or other solid structures, giving a natural adherence to the physical structure of a building.

This benefit of using physical walls as a feature is also captured in ultrasound (us) solutions, such as Dolphin[30], The Bat[31], or the Sonitor Indoor Positioning System[32]. These systems rely on having tags on the monitored objects transmitting an ID using ultrasound-waves which are captured by an ultrasound-microphone or even array of us-microphones (giving higher resolutions). Ultrasound, just like IR, does not permeate walls and is also attenuated rather quickly in air, which can be exploited to create several proximity-zones in one room. The main drawback over RF is the low bandwidth and slow wave propagation which limits the number of tags and amount of information that can be transmitted.

As with any attempt at classification of systems there are some that are hard to place in one category or the other, for instance systems such as the Cricket Indoor Location system[33]. This particular system relies on both transmissions using both RF and ultrasound and timing differences between the two forms of transmission.

3.4.5. Hybrid systems

While there are working examples based on several different types of technology, none are without inherent problems. For instance ultrasound transmission rates are low while RF-transmissions bleed through walls and cause interference.

There are solutions that attempt to mitigate the shortcomings caused by the individual transmission channels by using more than one[11] channel of communication. One such example is the Cricket Indoor Location System[33]. It is still a research project and thus not commercially available. The purpose of the multi-channel approach is often working around the short-comings of RF transmissions and ultrasound by using a combination of several forms to obtain location by for instance using Time-Difference-of-Arrival (TDOA)[12] between the ultrasound and RF signal.

Using TDOA the Cricket system obtains results in the range of an accuracy of 5 cm. Cricket has a position estimation accuracy of 10 cm and an orientation accuracy of 3 degrees[34]. The infrastructure required for Cricket is extensive, but according to the authors still relatively cheap (one of their design goals was for the “crickets” to cost less than $10). Similarly there are examples of systems that use a combination of RF and IR technology to solve the same challenges.

3.4.6. Alternative approaches

There is also some novel and creative use of existing technology to create location-based systems. These include video image recognition systems that can track objects moving in

---

[11] While several GPS solutions also rely on both microwave and traditional radio, the use of radio in GPS is to convey local correctional instructions rather than the location information.

[12] In

video streams (often already available in digital format from CCTV surveillance systems). This also includes systems that perform facial/feature recognition on images, thus being able to pinpoint not only an object, but identify it as one of a possibly large set of known objects.

Other sources of data can also be used for logical inference upon the location of objects, for instance using a swiped access card as a fact indicating that the owner is inside the controlled room, a positive identification of the user on video lends additional proof to the accuracy of the detection. In this group there is often a reliance on existing infrastructure already providing a service and the extraction of location information is often a secondary purpose or novelty feature of an existing system.

3.5. Summary

As described in the beginning of this chapter, location is a diverse concept that comes in a wide variety of forms. Simple queries about the location of an entity can produce answers from both the temporal and spatial domains as well as in terms of tasks or actions.

The diversity in the concept is similarly reflected in the diversity of the various examples of location systems identified in the latter part of the chapter. From this brief examination it is evident that no system would fit all purposes and the different systems obviously have different capabilities, strengths and weaknesses. However, we also saw that there are some fundamental algorithms and methods such as angulation and lateration that are employed regardless of transmission medium or where the location is computationally calculated.

4. Sonitor Indoor Positioning System

In Chapter 3 we reviewed several examples of available location system alternatives based on various methods and infrastructures.

In this thesis the Sonitor Indoor Positioning System was used for testing and experimenting with a location system. Sonitor Technologies ASA participates as a partner in the COSTT-project and the equipment was acquired through this relationship.

The chapter begins with an overall description of the Sonitor system, which parts it contains and how these interact. It then goes into more details on the individual parts and describing how the system actually works. The last two sections is devoted to a quick summary of how the system was installed in the laboratory and a summary of other installations.

4.1. Description

The Sonitor IPS is an ultrasound-based location system. Sonitor senses proximity via a detector capturing an ultrasound pulse emitted by a tag. The system then does a simple determination of which detector receives the signal, and uses that as a sign of proximity. In the event that several detectors should receive the same signal a “winner” is negotiated. This is usually[13] the detector with the stronger signal.

---

[13] Algorithm described as *usually* as there is some rudimentary filtering to smooth over tags that skip from one detector to another and back again in a relatively short period of time.

A Sonitor IPS consists of three parts; a server (software), detectors and base stations (hardware) and tags (hardware), as well as Ethernet connecting the stations to the server (using TCP/IP so it can be shared with existing office-infrastructure). In line with the ISO/IEC definition of RTLS, the server, detectors and base stations make up the location system infrastructure.

The server software is a TCP/IP-service that connects to the base stations specified in its configuration and collects data from the base stations. The server also have a TCP/IP interface where 3rd-party software clients can connect to receive the collated input (in practice an ASCII stream with mostly comma-separated data divided into the rudimentary location model[14] defined in the server configuration).

4.1.1. Server Software

The server software can be run on most common operating systems and standard commodity hardware. The server also supports multiple clients, thus moving a potential scalability issue from the specialised hardware over to commodity hardware. The base stations does in practice only support one system polling them at a time, while the server software running in a much larger environment scales much better.

Sonitor IPS Server version 3.3 (release date: 17th November 2008) were used through the work with this Thesis.

---

[14] The model is in fact transferred at connection time to the client ensuring a shared data model across server and clients.

4.1.2. Base stations

Figure 2 Sonitor Base Station (right) and a Detector (left)

There are several types of base stations, some with detectors built into them as well as a choice between wired Ethernet (the one in Figure 2 is a wired-network without a detector) and wireless network connection. The base stations have inputs for connecting several detectors as well as network. They can also be configured using a special software utility via the network to alter sensitivity and add scaling factors, as well as gathering some simple statistics, network configuration, and rebooting.

4.1.3. Detectors

The detectors (see Figure 2) are wired to the base stations and function as the systems ears picking up the transmissions from the tags. Simplified, the base stations are ultra sound microphones. There are also “wave-guides” available for mounting on the detectors to alter the profile of their coverage area. Equipped with a wave guide it is possible to create focused beam “hot spots” inside rooms already covered with narrower footprints than what the unguided detector can do.

4.1.4. Tags

Figure 3 Various Sonitor Tags. E-Tag (inset), two P-Tag cores and two P-Tag shells (open and closed)

Sonitor has currently two lines of tags, Patient (P-Tags) and Equipment (E-Tags) Tags, all shown in Figure 3. There are some minor variations between each line. On an abstract level the tags are functionally equivalent and there are only minor, but significant, differences between them. From a system perspective they are interchangeable in the sense that you cannot reliably distinguish them based on their performance from a system perspective.

The obvious difference apart from the physical design of the two tags is that the E-tags are equipped with two buttons. When pressing the buttons the tag transmits instantly and the data transmitted to the server indicates which button is pressed (A or B). The fact that the tag transmits instantly (and perhaps just as importantly, continuously while the button is held down) can be exploited to achieve near real-time functionality when required, but at the same time maintain less-frequent transmissions at other times to both conserve battery as well as avoiding medium contention[15].

The E-tags are also equipped with motion sensing capability which allows them to enter a state of sleep when they have remained stationary for a period of time (measured in multiples of their moving-state rate of transmission, e.g. indicated as 5x – which means that the tag will transmit 5 times after coming to rest). Whether or not the tag is active, is also

---

[15] A continuously transmitting tag will effectively block any other tags in the same area from transmitting.

transmitted along with the data from the tag. A single bit in the data transmitted indicates whether the tag is in motion or not, based on the motion sensing capability.

The P-tags on the other hand, does not include any motion sensing capability and thus transmit at regular intervals regardless of whether it has moved or not. This also allows the system to discern between a tag that has gone to sleep and a tag that has left the covered area. The motion sensing capability means that at comparable transmission intervals the E-Tags are more battery efficient, saving battery when possible.

The P-tags are designed to be worn clipped on bracelets or attached to a patient in another fashion. The tag consists of a disposable outer shell (as seen in the picture) with batteries and an inner core containing all the electronics. This allows the shell to be disposed when used to comply with hospital hygiene standards. The outer shell is also “splash proof” and designed to be hard to open without the correct tools (so the patients cannot remove or open it neither easily nor unintentionally).

4.2. Installation, Tuning and Calibration

Installation of the equipment is simple and in terms of technology is limited to running cables to the various devices (using RJ-45) and then some minor network configuration to allow server and base receivers to communicate.

After installation the location system should be tested and tuned to remove any artefacts created by the location it is installed in. This can be done by imposing a gain factor on a detector to make it more dominant in competition with others. This is especially useful in situations where there are several detectors with overlapping fields of coverage (for instance with several detectors in one room).

Tuning is an open-ended task and it has to be adapted not only to the physical features of the location the system is deployed into, but it also has to take into account movement patterns. The physical boundaries of a room is only partly the defining patterns of human movement, furniture and the function of the room can also impose important characteristics on how people and equipment move through a room and in turn define which areas need coverage and “dead zones” where coverage is less or not even required at all.

The installation of the system used in this thesis was done under guidance by a technician from Sonitor.

The equipment installed and available for the thesis work consisted of:

- 3 DBAS-B-02 Base receivers.
- 10 DSat-B-01 Satellite Receivers (detectors, or ultrasound microphones)
- 2 DSat-B-02D Satellite Receivers with wave guide (detectors, or ultrasound microphones)
- 10 E-Tag ultrasound emitting equipment tags (10 second transmission interval)
- 10 E-Tag ultrasound emitting equipment tags (30 second transmission interval)
- 10 P-Tag ultrasound emitting equipment tags (5 second transmission interval)
- 10 P-Tag ultrasound emitting equipment tags (10 second transmission interval)
- Sonitor IPS Server software v3.3, and setup utilities as well as access to the Sonitor Partner Website for manuals, etc.

4.3. Sample deployments

Sonitor IPS is currently not deployed in any Norwegian hospitals. However their website[16] lists, at the time of writing, deployments in 14 American hospitals ranging from hospital-wide installations to more limited tracking of key personnel or equipment.

Other notable, but temporary, installations involve the project “FindMyFriends” as deployed during the student festival “UKA 2007”:

> *Approximately 3.000 students will wear a Sonitor® Tag when they enter the famous student society building "Samfundet" (The Society). The students’ positions will be known through a "Facebook"-like internet application called "FindMyFriends". Profile information for each student like sex, civil status, and field of study can be displayed. “Samfundet” is a large and complex building and with "FindMyFriends" you can check out which of your friends are inside - and where. The application can be accessed from home or on one of the stations inside “Samfundet”.*

-Sonitor Press release, August 23rd 2007.

---

[16] http://www.sonitor.com/ Visited 01.03.2009

5. Operational Capabilities

This chapter investigates how various characteristics related to location systems reviewed in Chapter 3 have to be taken into account when assessing a location system with the prospect of installing it into a hospital. This involves trying to uncover how these various parameters might affect the operation of an installed system.

The chapter begins with a discussion about operational capabilities and the various dimensions that describe these capabilities. Using the operational capability concept relevant literature is search for examples of location system use in the hospital domain. The examples are then classified according to the identified dimensions.

5.1. What are Operational Capabilities

The Oxford English Dictionary defines *Capable* as "having the needful capacity, power, or fitness for (some specified purpose or activity)". As discussed in Chapters 3 and to some extent in Chapter 4 there are numerous properties of location systems in general that can be explored and to some extent objectively measured and quantified. These characteristics play different roles depending on the features of the system. As input for an assessment however, these numbers are of little meaning taken out of context.

So rather than focusing on particular characteristics, which vary between systems, we have opted to look at what capabilities (i.e. the fitness for some specified purpose) does a system need to have. Instead of describing the capability requirements through defining numerical ranges for various characteristics, we have chosen to identify these domain-specific purposes or activities that might be necessary to be support in a hospital.

5.2. Dimensions

To closer examine these purposes we have identified three dimensions in which they can be classified. The dimensions identified are: granularity, resolution and concurrency.

5.2.1. Granularity and Amounts of Infrastructure

Almost regardless of the location system technology and its characteristics, there is some infrastructure that needs to be in place for it to operate. Even with systems relying on infrastructure intended for other purposes, it might be necessary to extend or alter the existing infrastructural layout to provide sufficient coverage. So when trying to specify and design a new system it is necessary to make some decisions about where and how much infrastructure needs to be installed or improved. This means having to make decisions about which floors, spaces within floors or particular rooms need coverage to detect enough data to be able to inform a computer system with both reliable and meaningful data.

Table 1 show a possible scale of granularity and infrastructure amounts, which will be used to analyse the operational capabilities sought to assess.

Low ↔ High

Sporadic Proximity (“choke-points”)	Transport spaces	Strategic coverage	Functional rooms	Every room	Strategic zoning	Zones every-where	Micro-Zones[17]
Low Granularity		Medium				High Granularity	

Table 1 Scale of granularity and infrastructure

This is of course not a set of discrete steps, but rather a continuous scale from a single proximity detection at one extreme to a level of granularity able to detect virtually any change at all (sometimes referred to as sub-person granularity). With micro-zones or sub-person granularity the system will be able to detect location down to centimetres or even finer.

---

[17] Micro-Zones are used to denote zones so small they are approaching centimetre or even better granularity.

Infrastructural Granularity and System Granularity

In the identification of capability and the granularity requirements the assessment does not, on purpose, make a clear distinction between what is caused by granularity and what is caused by the size of the infrastructure. The reason for this indistinctness is the somewhat complex relationship between these two factors. Depending on the system, there is often a strong interdependency between infrastructure and granularity. Higher amounts of infrastructure often improve the granularity of the system. However, the granularity of a system often has a fixed upper limit based on the physics of the system.

So rather than favouring certain types of systems where the presence infrastructure implies a fixed granularity over systems where the differing amounts of infrastructure affects granularity, the granularity notion is used to cover both situations.

For systems with fixed granularity, it is usually a matter of identifying interesting rooms. On the other hand, for systems with varying granularity parts of the assessment is finding the right amount of infrastructure where one can achieve the granularity required, but on the other hand not needlessly limiting the granularity by not installing enough infrastructure. For these systems there usually is a threshold where increase in infrastructure will not yield a similar (or even any) increase in granularity. Hence it is important to understand where one needs the maximum granularity and where lower levels are acceptable in order to be able to predict how much infrastructure to deploy.

Example of Increasing Granularity

When using a location system to detect whether orderlies are available (i.e. idle) there are solutions of various levels of granularity.

At one extreme, with a single proximity sensor, one could detect when the orderlies are in the proximity of their station. However, one could easily find examples where this type of idle-detection would fail. For instance the orderlies might want to get coffee when they are idle and hence capitalising on their idle state to do unimportant and interruptible errand they become out of reach for the simple proximity and would be considered, erroneously, as busy in the eyes of the system.

Stepping along the scale in Table 1, the similar case could be solved by introducing proximity detectors in all areas where orderlies should be considered idle. This could for instance be around a coffee-machine, toilets[18], break rooms, etc. This increase in the amount of infrastructure is leading to a more reliable detection of the status of the orderlies.

Jumping to the other extreme, one could envision a system that is able to track their movement within rooms so that their status could be changed to idle when they hang around any space for a period of time or even more intelligently consider the type of space and then apply different schemes of identifying idleness based on this space.

5.2.2. Resolution

Closely related to granularity is the temporal resolution of the system. The resolution is an indication of how long time a system needs to detect a change of location. This obviously has an effect on the capabilities of the system. Operating with a low resolution (i.e. long time to detection) there is an increasing amount of information that is not captured between the updates of the system. Conversely there is a threshold that when exceeded, the extra location data does not offer additional information over the data collected at lower resolutions.

The required resolution is dependent on what is being tracked with the system, i.e. the speed of which things move is the restricting factor. The faster objects move, the higher resolution might be needed to meaningfully track them. For indoor systems in hospitals there is a practical maximum velocity of any object, usually closely related to the speed of which humans move. At the other end, there is no lower bound for resolution, but at some point the system will cease to give reliable answers as the objects will have moved before the next update. Thus the uncertainty related to the detected location will become unacceptable.

The walking speed[19] of humans are taken to be on average 5 kmh^{-1}, which means that in one second a human can walk between 1 to 1.5-metres. This gives an indication of where the

---

[18] Toilets and wardrobes are usually sensitive areas with respect to privacy, but for the sake of the example let the availability monitor not disclose their whereabouts, just their status as “idle” or “busy”.

[19] The speed given in literature varies from 3 kmh^{-1} to 7 kmh^{-1} depending on various factors.

level for medium resolution lies. For the purpose of assessment a table based on the average speed is proposed in Table 2.

Low ⟷ High

Minutes	Seconds	Sub-second
Low Resolution	Medium	High Resolution

Table 2 Scale of Resolution

Similarly to the scale for granularity, there are no absolute or discrete levels of resolution of which to choose from. Some capabilities will have clear thresholds in terms of resolution to achieve a reliable result, while other is more fleeting in terms resolution requirements.

5.2.3. Concurrency

The third dimension is concurrency. Concurrency is related to the other two dimensions and depending on the technology used they influence each other in different ways. Unlike the other two it is easier to directly control. The highest level of concurrency is directly related to the number of tags that are deployed. Because of this it is not included in the operational capability matrix below.

While as the other dimensions scale almost independently of the number of capabilities, concurrency will grow almost linearly with the number of supported capabilities. That is to say that each type of objects tracked will add a certain amount to the required level of concurrency to operate. Similarly it is possible for most systems to find theoretical or practical maximums for concurrency and compare this to the expected number of tags in any area. Depending on the system and how concurrency relates to the other dimensions is it possible to adjust the granularity to mitigate concurrency issues or evaluate whether loss in resolution due to concurrency is acceptable.

5.3. Operational Capabilities

The operational capabilities identified are based on the use of IPS in similar situations found in literature. The references indicate the source.

5.3.1. Method

The source of the use cases are mainly extracted from literature describing use and/or evaluations of IPS applications within the domain. The searches were performed in PubMed[20] between late June and early July 2009. The search-phrases used were permutations of the following terms: IPS, indoor, positioning, location, locating, and tracking. The titles and abstracts were browsed to identify appropriate papers. The papers were then consulted for the characteristics of the challenge attempted solved (or described). The references of identified papers were also review in search of candidates for inclusion.

The extraction of intent or purpose from the various identified papers involves a certain amount of subjective interpretation. For instance there is a fine distinction between pure inventory systems and location systems. While the technology used might be the same, there is certainly a distinction in the use of the gathered data. There is a growing amount of literature discussing the introduction and use of RFID in the healthcare sector. Many of these publications are case based and focus on the functionality realised rather than the technology used.

This is by no means intended to be an exhaustive exploration of hospital implementations, but rather a meaningful sample or set of challenges that can be answered with a location system.

---

[20] http://www.ncbi.nlm.nih.gov/pubmed/ accessed 01.07.2009

PubMed/Entrez is a search engine for searching the Medical Literature Analysis and Retrieval System (MEDLINE) which indexes over 5000 publications on life-sciences and biomedical topics. Entrez also includes approximately another 20 databases on health sciences.

5.3.2. Findings

Operational Capabilities		Granularity			Resolution		
		Low	Medium	High	Low	Medium	High
1	Find a particular member of staff		X			X	
2	Determining the location of an in-patient[35, 36]			X		X	
3	Find a particular out-patient[36]		X			X	
4	Find a piece of equipment (IV-pump, Ultrasound machine)[6, 36]		X		X		
5	Locating the nearest staff to a particular room		X				X
6	Determining availability of on-call physicians		X		X		
7	Determining the availability of beds[35]			X	X		
8	Indicating availability of in demand rooms	X			X		
9	Approximate time of arrival of a paged staff member	X				X	
10	Being able to indicate progression status for a pre-defined patient trajectory[7]	X			X		
11	Auto Log on/off hospital computer systems[37]			X			X

Operational Capabilities		Granularity			Resolution		
		Low	*Medium*	*High*	*Low*	*Medium*	*High*
12	Measuring waiting time for patients[38]	X			X		
13	Monitor the utilisation of a piece of equipment[39]		X		X		
14	Measuring complete turn-around time for a patient[7]	X			X		
15	Detecting the progression status of examination with well-defined trajectories[40]			X	X		
16	Tracking hygiene regulations compliance[37]		X			X	
17	Schedule equipment maintenance based on location and utilisation[41]		X		X		
18	Detecting whether equipment has been sterilized since use[42]	X			X		
19	Improve hospital security by notifying patients/employees in restricted areas[41]	X				X	
20	Detecting patients that has fallen[37]			X	X		

Table 3 Operational capabilities against granularity requirements

Tallying up the granularity and resolution requirements yields the matrix of results shown in Table 4.

Resolution	High		5,	11,
	Medium	9,19	1,3,16,	2,
	Low	8,10,12,14,18,	4,6,13,17,	7,15,20
		Low	Medium	High
		Granularity		

Table 4 Matrix of Granularity and Resolution from Table 3

There are few systems that would span the whole range in a sensible way. Some IPS technologies are only capable of reliably solving low-granularity tasks. For instance is several of the RFID-based systems available are in general better suited towards tasks such as inventory control and proximity rather than large scale high-granularity positioning for people and equipment. Correspondingly there are technologies such as several of the hybrid systems that would be under-utilised or inappropriately applied if only seeking to cover low-granularity applications.

5.4. Operational Capabilities as an Assessment Tool

The purpose of mapping out the operational capabilities sought in a location system is to better understand what one *should* assess when comparing location systems. Exclusively comparing the absolute values of granularity, resolution and accuracy will not uncover how suitable a particular system is for the job at hand, but rather give a comparison of the systems to each other. To be able to really assess the suitability of a location system for a particular task, one has to understand the relationship between the capabilities required by the task, and the relationship between these capabilities and the location system characteristics (see section 3.2 in Chapter 3).

Using these dimensions it is possible to design experiments to attempt to find ranges of the dimensions supported by particular location system (shown through experimentation). It is then possible to map the dimensions supported by the location system back onto the matrix of operational capabilities. This creates a link between the technical characteristics of a location system and the descriptions of capabilities of a location system.

At an early stage of a procurement process, a list of operational capabilities might also be able to offer valuable input towards what reasonably can be expected of a location system, as well as indicate to whether or not the capabilities sought represent challenges that are difficult to solve.

5.5. The Need for Additional Dimensions

In this Thesis we chose to use only two dimensions to categorise the various capabilities for the particular domain. While this gives a rough, but useful categorisation of demands into 9 quadrants, one could easily add additional dimensions.

The identification of suitable dimensions is hard. To find dimensions abstract enough to be applicable across various location system technologies, but at the same be both understandable and valuable as a metric in an assessment. There is also a necessity for the dimensions to be reasonably objective and testable.

The reason for adapting only two dimensions is the fact that a more detailed classification increases the difficulty in allocating dimensional values to the capabilities without really adding too much additional value to the matrix. Similarly it might be useful to re-categorise the capabilities in more fine-grained steps along the existing dimensions creating a finer mesh. However this is perhaps more useful if the point is to distinguish between similar systems where the differences in capabilities are.

5.6. Summary

In this Chapter we have described what operational capabilities are and which dimensions that can be used to describe them.

Using location system-related search terms in PubMed, we identified 20 different purposes that location systems are used for in hospitals. These 20 purposes where then classified in terms of the identified dimensions to form a matrix of example purposes. This matrix can be used to describe the operational capabilities required for a location system in a hospital setting.

6. Location System Testing

Chapter 5 identified three dimensions that can be used to describe the capability required to serve a particular purpose. Using the location system described in Chapter 4 this chapter assesses how values these dimensions can be found for the Sonitor system.

The chapter begins with a description of the materials used for experiments and the underlying theory behind the experiment design. Section 6.3 and 6.4 describes the individual tests in detail. Section 6.5 is an overall evaluation of test results in terms of the capability of the Sonitor system.

6.1. Laboratory

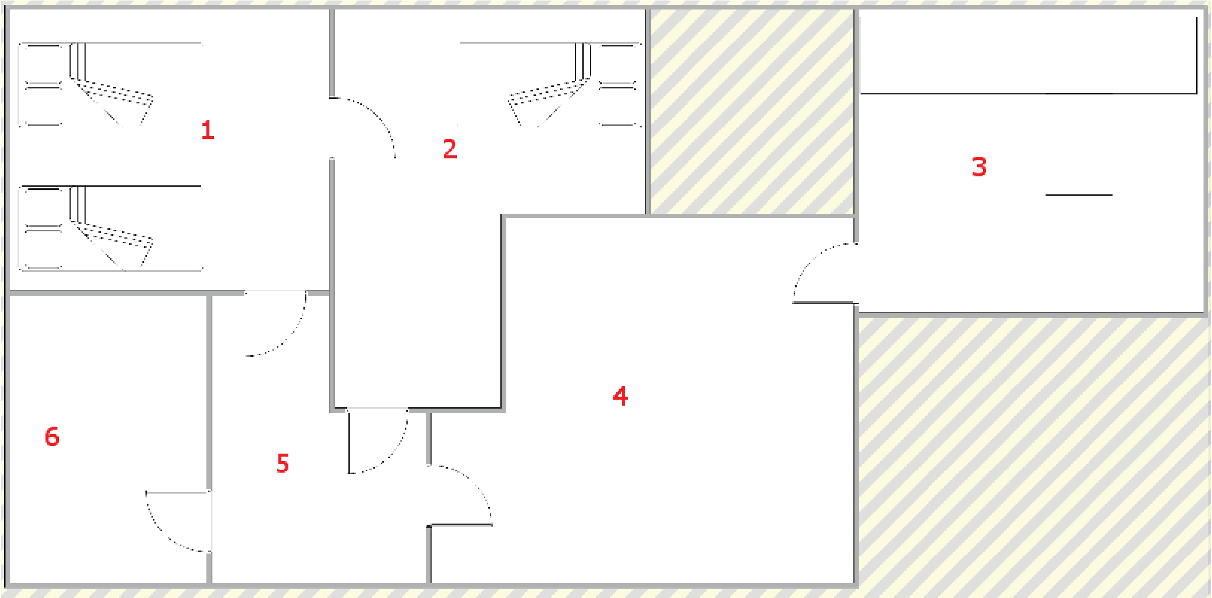


Figure 4 Floor plan of the NSEP Usability Laboratory

The location system infrastructure from Sonitor Technologies was installed in the usability laboratory located at The Norwegian Electronic Health Record Research Centre (NSEP). This is a laboratory designed and equipped to perform a large variety of usability-related tests and experiments.

The lab includes a test environment with reconfigurable walls, which enables the simulation of different environments. It is equipped with remote-controlled audio- and video-recording equipment to provide documentation of the tests in real-time. All this equipment is controlled from an adjoining control room. The control room also contains large screens to enable test controllers and technicians to watch and control the tests in real time from the comfort of the control room, not impending on any test candidates.

The laboratory is equipped with hospital beds and other furniture commonly found in hospitals to make it appear as a small hospital department during a test (see Table 5 for full listing of rooms, room numbers referring to Figure 4).

Room number	Symbolic room name	Number of detectors
1	"Patient room 1"	3
2	"Patient room 2"	2 (1 with wave guide)
3	"Control Room"	1
4	"Lobby"	2
5	"Corridor"	2
6	"Office"	1

Table 5 Room list

6.2. Experiment Design

As initial steps towards high-level functional experiments it is necessary to test the basic fundamental properties of the system. This ensures that any functional experiment is designed in accordance with the fundamental properties of the system. It would also enable to distinguish between the effects of the environment and the features of the system. To accomplish this, the experiments were set up as a loosely connected series experiments, each experiment with its own aim and hypothesis, but as a piece of a larger picture.

The main factor affecting the reproducibility and hence the scientific method and value of these results is the environment the tests were performed within. To guarantee reproducibility and to control the limitations of the results the experiments would have to be performed in a clean controlled environment. This would make it possible to, with a high degree of certainty, be able to distinguish environmental influences from the system performance.

However, in this investigation emphasis has been put upon performance in realistic settings for use in health care. This is why the tests have been performed in the usability laboratory at NSEP in a reasonably realistic hospital environment. This realistic, but uncontrolled, environment does however mean that any side effects and unwanted or adverse influences cannot be directly controlled, but rather mitigated through use of statistics and higher number of repetitions.

The collected data were analysed using Microsoft Excel 2007, the statistical software Minitab® 15.1.30.0 as well as some purpose-made minor scripts and utilities (see Appendix C for details).

6.3. Baseline functional tests

The first round of testing was designed to obtain fundamental properties with the system that both guide and impose limitations upon both later tests as well as functionality in later system development and implementation. The overall idea was to devise a set of basic and rather simple tests that collectively can inform of basic capabilities, or restrictions, of the location system.

6.3.1. Individual tag strength

Rudimentary experimentation showed that there was some variation in the strength of the pulses received from the various tags. This was expected due to inherent variability in any production process (and electronics). The rationale for this experiment was to establish a set of statistical parameters for interpreting later results.

The strength of pulses between individual tags are of low importance in a functioning system, as the current incarnation of the Sonitor system does not use this information and by default it is consumed and abstracted away at the lowest possible level. However the individual variation (or at least a statistical prediction of it) is interesting and important for ensuring reproducibility and for interpreting the results of baseline tests. It will also uncover whether or not this is a metric that can be used for inferring information about the located tag.

Hypothesis

H1. The strength of pulses is consistent over time.

H2. The strength of pulses is consistent across tags (given equal distance and detector)

Equipment

- Single Detector
- Different set of tags (several of each type for comparison)

Method

Using a single detector to ensure reproducibility across tags, a number of pulses are to be collected over time. Then using the log files the strength of the individual pulses can be extracted and compared both for consistency across time (as per H1) and across tags (as per H2).

Results

In practice there were 4 different types of tags, two types of P-Tags (the difference being the transmission interval, 10-seconds and 30-seconds, respectively) and two types of E-Tags (difference being transmission interval, 5-second and 7-second as well as the number of resting transmissions[21]). This means that as well as comparing the individual tags within one type, it is also interesting to compare types.

H1. The strength of the pulses is consistent over time.

---

[21] i.e. the number of pulses the tag transmit when the motion detector is idle before falling to sleep.

Optimally the system would emit pulses of consistent strength over time so that the range (and other affected parameters) stays consistent regardless of external influences. Doing Trend Analysis on the data series produced by this test revealed no major cycles, patterns or trends. The results varied between the various tags; this can possibly be attributed to the fact that the tests were not run at the same time (they were run consecutively for the different types of tags) which means that any transient environmental effects could have affected one type of tag and not the other.

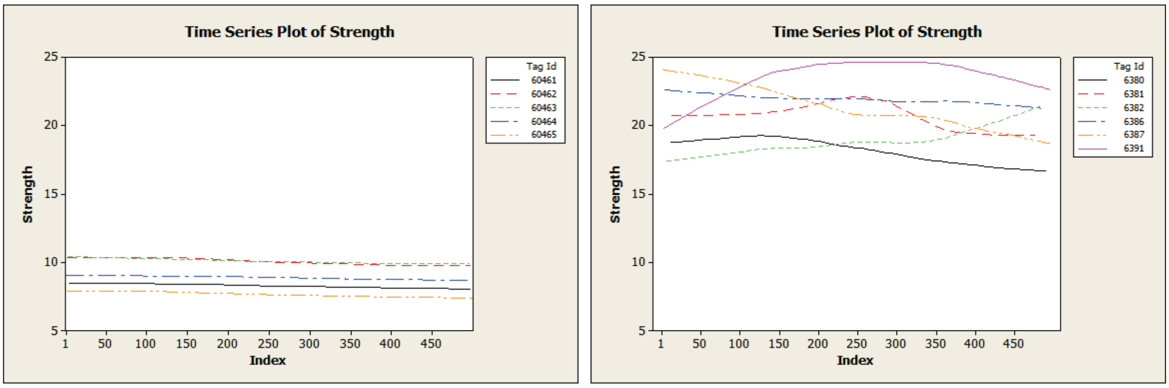


Figure 5 Smoothed (2 Step Average) Time Series Plot of 10-second P-Tags (Left) and 10-second E-Tags (right)

As we can see from the graphs in Figure 5, there seems to be some variation over time. This is most pronounced in the E-Tags (shown right). The P-Tags seems to be fairly stable, but on a longer perspective shows a minor, but observable tendency of decrease in strength. However, there is no obvious (and universal) pattern that appears to affect the functionality of the system on an overall level.

For the E-Tags it was a challenge to keep them transmitting (i.e. not falling to sleep) during the test which means they had to be moved ever so slightly throughout the test. It is not possible to say whether or not this has affected the results, and as such it would seem necessary to obtain more data before accepting or refusing the hypothesis. The currently available data suggests that there is cause for further investigation. Attempts to perform the same experiment with the 7-second E-Tags produced similar results, though with different curves. Not really shedding additional light on the hypothesis.

As mentioned, the P-Tags seemed to have little variation over the relatively short period of time (500 pulses is about a 15-minute time span) of the test.

H2. The strength of pulses is consistent across tags (given equal distance and detector)

The second hypothesis was formed to investigate whether the tags could be used interchangeably in testing (and production) and to inform later tests of possible ranges of what could be expected from the individual tags. The overall answer is that there is a significant difference in the strength of tags, particularly between P- and E- tags. For equivalent distances in the same setup, the equipment tags were received on average twice as strong as the patient tags.

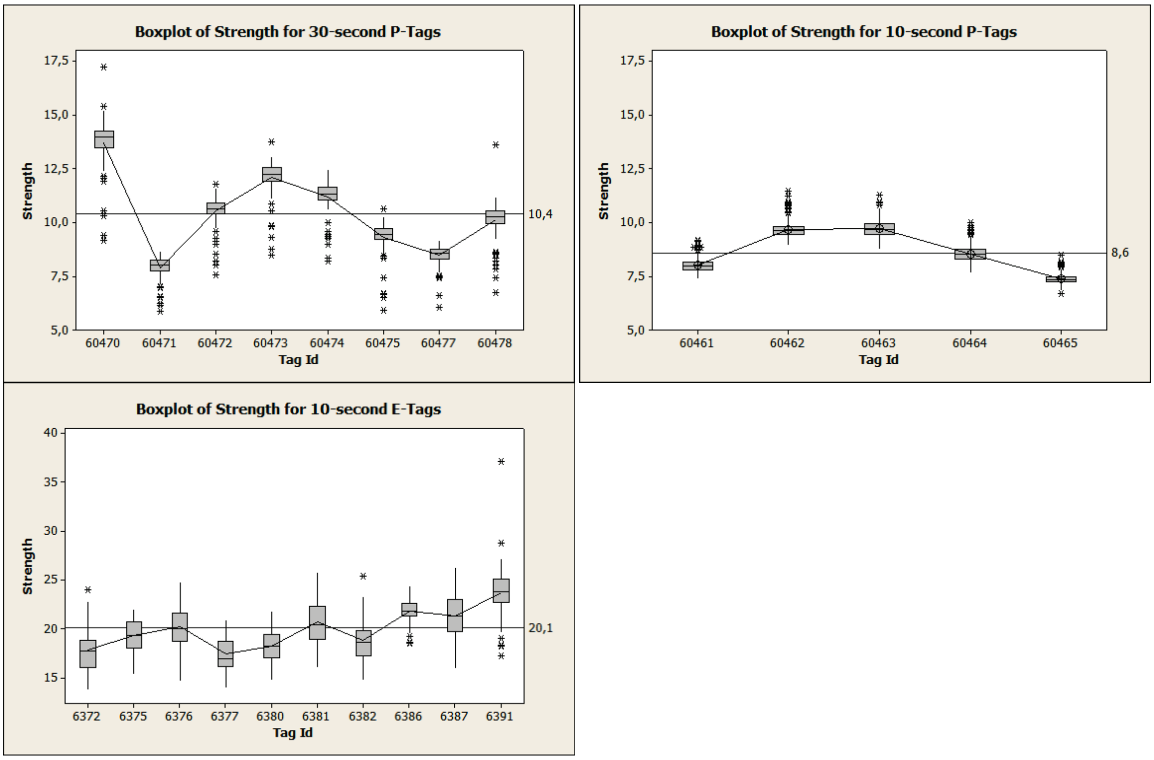


Figure 6 Box plots of various tag strengths[22]

Without enough data to support a conclusion but still a noteworthy observation, is the tendency for P-Tag outliers to be stronger than the average. The trend seems to be opposite for the E-Tags.

---

[22] The whiskers are correspondingly upper limit = $Q_3 + 1.5(Q_3-Q_1)$ and lower limit = $Q_1 - 1.5(Q_3-Q_1)$. Any outlier beyond these limits is denoted by a star.

Note the use of different Y-axis in the various graphs.

The descriptive statistics for these tests are included in Appendix A.

6.3.2. Detector Consistency

Variation between different tags could be observed in the preliminary work leading up to the experiment design. It was however harder to observe whether there was significant variation between detectors. Partly because it is masked by the variations in signal strength as well as the lack of identical rooms in the laboratory installation and as such there were no detector configurations that were equivalent in terms of room layout and orientation. This test is therefore aimed at obtaining comparable measurements from several detectors.

Hypothesis

H1. Two detectors with equivalent distance to tags of the same orientation will receive pulses of equal strengths

Equipment

- Two detectors
- 9 P-Tags

Method

Two detectors are mounted side-by-side (horizontally as seen from tags) and a line of nine tags are placed 2-metres from the detectors (with the fifth tag between the two detectors) and a number of pulses are to be collected.

The collected log files can then be analysed by comparing the detection of the same pulse by the two detectors.

Results

The results from the non-calibrated detectors yielded results hinting at potentially statistically significant differences between the detectors.

The initial idea behind having several tags was to both counter any effect from variations in a single tag (and thus obtaining meaningful averages rather than single pulses) as well as trying to even out the effect of the small difference in angle of arrival for the different tags. This latter effect was also closer inspected by dividing the tags in two groups and comparing the results.

Detector	N	Mean	Std. Dev	Minimum	Median	Maximum
A (#608040413)	509	7.4250	1.6816	5.6460	6.9270	23.5120
B (#608040410)	509	7.6933	1.7394	5.8740	7.6370	23.8440
C (#608040200)	506	10.200	2.681	7.209	10.031	29.810
D (#608040421)	506	15.299	3.714	10.747	14.236	37.002

Table 6 Descriptive Statistics for Detectors

As we can see from Table 6, there is a significant difference between the different receivers, actually to the point where both the mean and median of the receiver D is on the order of twice as strong as the values for detectors A and B.

6.3.3. Range and Angle of Arrival (AoA)

As with any other physical system, there are some physical properties relating to issues such as wave propagation through air, to more concrete design choices made in developing the system – all of which affect the range and coverage of an implementation. This set of tests is therefore aimed at gathering empirical data about the range of the system as well as the effect of the angle between the receiver and the tag (i.e. the angle of arrival).

There are interesting physical properties of the various detectors that can be exploited to create more fine-grained systems, for instance exploiting the coverage area of individual detectors to enable more advanced location schemes than mere proximity (for instance zones within larger rooms, etc.). These properties are for instance the variation in strength by distance or angle, as important for lateration and angulation methods of positioning.

Hypothesis

H1. The strength of detection is inverse proportional to the distance between tag and detector.

H2. The detector has a 180° field of reception and signals received beyond this angle is effectively attenuated[23] so to be discernable from signals received inside this angle.

Equipment

- One detector
- Tags

Method

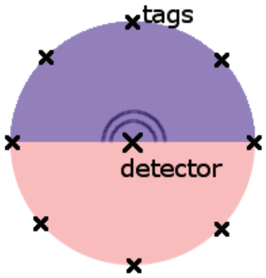


Figure 7 Tag placement illustrated as seen from above

The tags are to be placed as illustrated in Figure 7. The receiving-end of the detector is facing up in the illustration with the tags in the light red area in fact being behind the intended coverage area of the detector. The tags are to be placed as the spots indicated in the figure (at 0°, ±45°, ±90°, ±135° and 180°, with 0° being dead ahead as seen from the detector).

Results

The range of the system was expected to (as formulated in the first hypothesis for this particular test), due to fundamental physical properties, be modelled with strength being inverse proportional to distance between detector and tag.

---

[23] No specific measure of strength is given in the hypothesis due to the lack of any objective method of measuring it. Rather the signal strength as reported by the Sonitor server is used as an indication of attenuation.

To account for the uncontrolled variation in received signal strength several measurements were taken for each angle and distance. The time allowed for measurements of each angle was attempted to keep similar (to avoid unfair bias) which led to the result that the more extreme angles of arrival has fewer receptions. The time allotted was calculated to obtain n=10. See Appendix B for tables of descriptive statistics for this particular test (including number of pulses). In practice this mainly affected tags outside the 180° field put out in the hypothesis. For distances 100-centimetres and beyond, tags at -90° through 90° all managed a full 10 pulses.

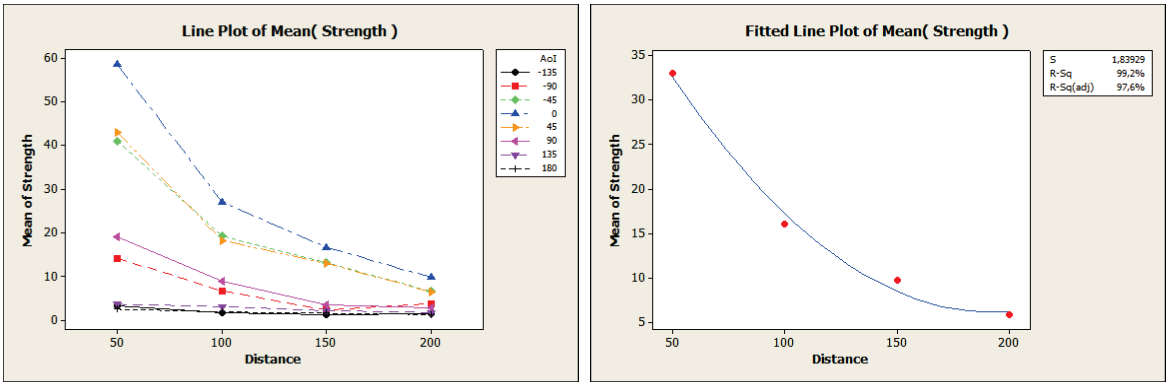


Figure 8 Line plots of Strength versus Distance (left), with a polynomial regression (right)

As we can see from the graphs in Figure 8, there is evidence to support hypothesis 1 – whether there is a relation between distance and strength. With albeit a small set of distances, a regression (see the rightmost plot in Figure 8) to a second degree polynomial seems to be a good predictor for the strength. This is in line with what would be expected from the physics of sound, where the inverse square law[24] is used to model sound intensity with respect to distance from the source.

---

[24] $I \propto \frac{1}{r^2}$

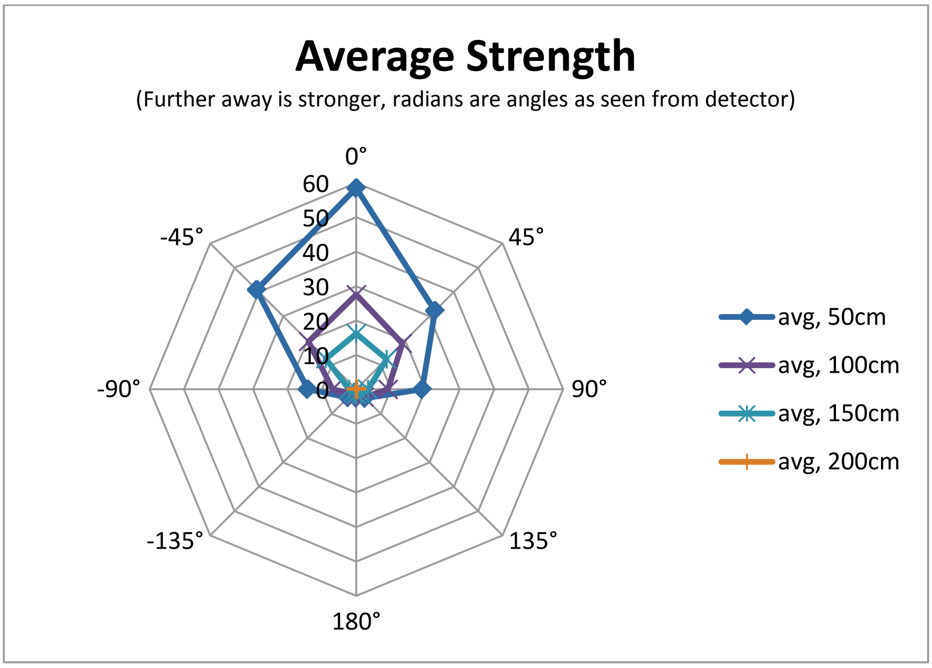


Figure 9 Radar-plot of strength versus angle and distance

The examination of the evidence for the second hypothesis, whether the angle of arrival is important, is illustrated in Figure 9. The figure is drawn with the detector seen from above and correspondingly 0° being dead ahead from the detectors perspective. We can see quite a clear distinction between tags at ±90° and those at ±135° and 180°. This is also evidence in support of the hypothesis. The figure also seems to show some degree of symmetry along the 0°-180° axis.

6.3.4. Tag Concurrency

The tags transmit their own identity (along with some status information) using ultra sonic sound waves. These are detected by an ultrasound microphone in a detector.

The tags transmit on the same frequency; hence they share a common transmission media (similar to other wired and wireless technologies relying on a shared media/bus, e.g. Wi-Fi, Ethernet, GSM, etc.). To avoid collisions between individual transmissions the individual tags use the Carrier Sense Multiple Access (CSMA) collision avoidance protocol. In practical terms, this means that each tag listens (sense) on the medium (carrier) and if the medium is busy it waits for a random period of time before sensing again, repeating until it is available for transmission.

The tags transmit 28 bits (their own id plus some status bits for indication of battery state, movement sensor and button states). The transmission rate is specified by Sonitor to be 50 bits per second, which means that each transmission should take 0.56 seconds to complete. Coupled with the delay caused by the carrier sensing, the complete time slot required is in practice closer to 0.7 seconds. Given this rather long time for the transmission of each pulse, there is a significant risk of saturating the medium by introducing enough rather rapidly transmitting tags into the same environment.

Rewriting this information into a formula we get:

$$tag\ update\ rate \times \frac{1}{\left(\frac{28\ bits\ (transmission\ length)}{50\,\frac{bits}{s}\,(transmission\ rate)}\right)}$$

Using this formula one can obtain theoretical numbers for the number of concurrent tags. We can see that for a tag update rate of 30 seconds this allows for over 50 tags in range of one detector before there is a scarcity of medium access. However, faster tags (or even a mix of tags) will quickly reduce this number. The 5-second tags will theoretically run out of medium already at 9 tags, and correspondingly just over 17 for the 10-second version.

When examining log files with complete detection logs, the timing between the pulses should be equal to the pulse-rate of the tags. It should also show regularity at this interval, i.e. a 30-second tag should be received every 30 seconds (including some buffer for collision avoidance with other tags). This Inter-Arrival Time (IAT) is a good indicator for whether the area around a detector is saturated with tags. IAT is found by calculating the difference between the individual timestamps of the pulses.

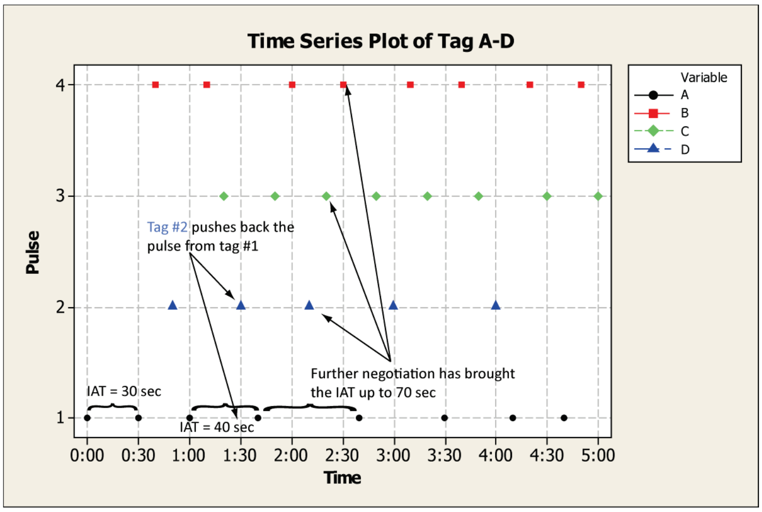


Figure 10 Illustration showing the effects of CSMA on IAT

Figure 10 illustrates the effect that collision avoidance has on the IAT of an individual tag. The illustration is a hypothetical situation where each tag requires a 10-second slot. At the introduction of the fourth tag (Tag#3) at T=1:20 there is not enough medium access to sustain four tags. At T=1:30 when Tag#1 is supposed to transmit, it is pushed back by the transmission by Tag#2 (which already lost the negotiation with Tag#4 10 seconds earlier). This lost negotiation for Tag#1 brings the IAT between its third and fourth transmission up to 40-seconds. At the next transmission it is “unlucky” and looses to three more tags and the IAT is brought up to 70-seconds. At this point Tag#1 is effectively starved of access to the transmission medium.

Hypothesis

H1. The Inter-Arrival Time (IAT) of the detections will be stable until saturation point is reached.

Equipment

- 1 Detector
- P-Tags

Method

All tags are registered in the system. A working infrastructure with one detector is started and the server is set to log all detections to file. The 10-second pulse-rate P-Tags will be used for this test. This brings the required number of tags down to a manageable number, as well as more predictable behaviour because of the lack of motion detection.

The server will be started in order for logging to commence. The e-tags will then be turned on one-by-one in >30 second intervals. The log-file will then be analysed and the time between detections (IAT) for each tag will be extracted.

Results

The tag concurrency experiment was run twice, once with a two minute interval between the introductions of new tags, and based on the results repeated with a five minute interval.

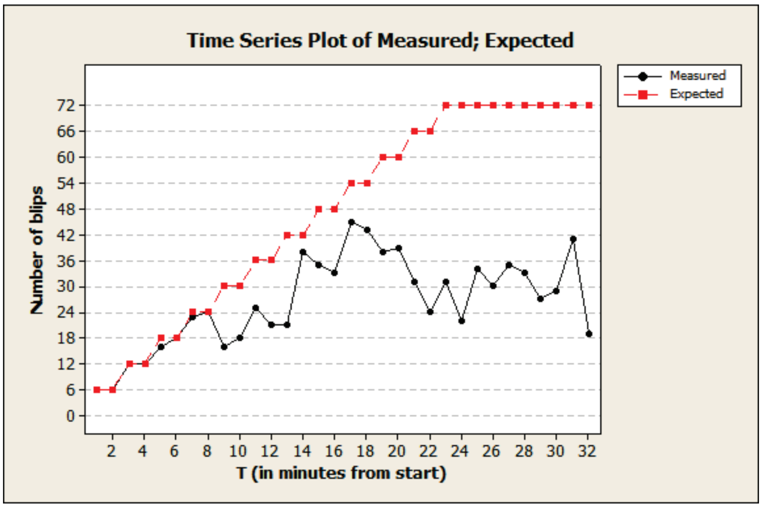


Figure 11 Expected number of detections versus measured detections for 2-minute

As seen in Figure 11, the rapid introduction of new tags compared to the relatively (as opposed to introduction rate) low pulse-rate of the tags (30-seconds) the system faced difficulties maintaining reliable detection of all tags after the introduction of the 5th tag. After the introduction of the 9th tag (well beyond the calculated saturation of the medium) the system performance seems to degrade even further.

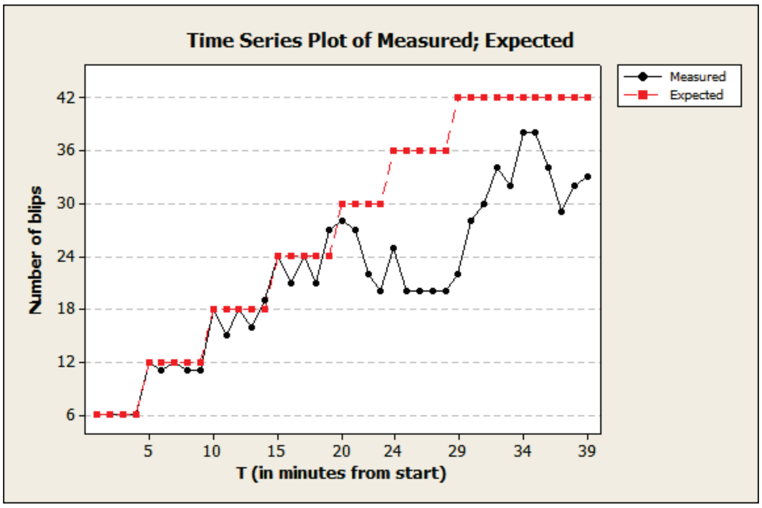


Figure 12 Expected number of detections versus measured detections for 5-minute

Repeating the same experiment with a longer interval between introduction of new tags (5-minute rather than 2-minute), yielded the graph shown in Figure 12. The overall findings are similar to those of the 2-minute experiment; the threshold for system degradation seems to be at the introduction of tag number 5. In this experiment, only seven tags were introduced to the system to give it a fair chance of negotiating a steady state (with seven tags being within the range indicated by the formula given in the test description). Even after 10 minutes, the system had not negotiated itself into a steady state.

Given a closer look at the inter-arrival times (see section 6.35.4 for a description of IAT as a metric) for the individual tags, there are sections in both tests where individual tags are starved for medium access and thus cannot be located due to lack of medium access, i.e. it is impossible to distinguish between a starved tag and a tag that has left the area covered by the detector.

As seen in Figure 13, there are periods in the test where one tag is starved for prolonged periods of time, for instance Tag 60463 from T=1200 to T=1440, and from T=1140 there are only single unconnected series of pulses detected. The time to discovery (the time from the tag is introduced to it is recognised by the detector) is also increased for several of the latter tags.

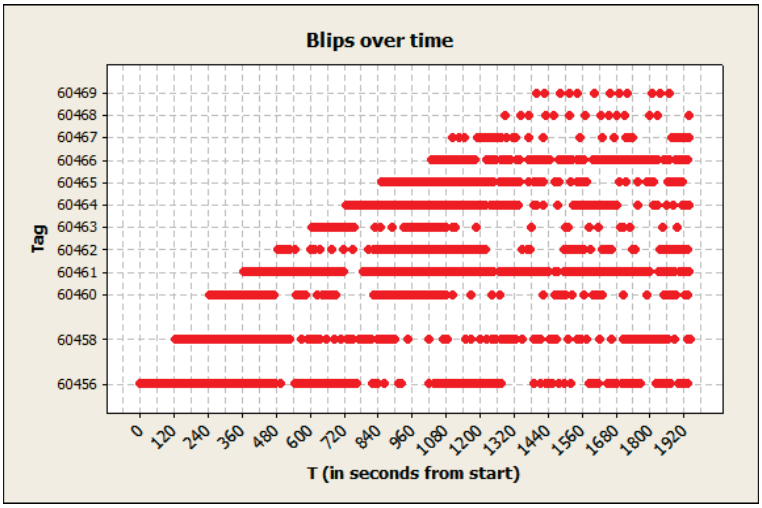


Figure 13 Pulse plot

6.4. More Functional Tests

The next set of tests is designed closer t directly testing the capabilities sought investigated. Having established some basis for interpretation and knowledge in the previous set of test, this section includes testing information quality parameters as defined in section 3.2.3.

6.4.1. Room-Level Precision and Accuracy

One of the common reasons to choose ultrasound over RF-based technology is the strength of ultrasound when it comes to adhering to physical boundaries – namely room-level precision. Whilst a RF-based system would have problems with signals escaping through walls and doorways, ultrasound waves should in theory be effectively stopped by such obstacles due to basic differences between sound and electromagnetic waves.

Hypothesis

H1. Ultrasound is effectively blocked by walls, so with >99% accuracy >99% of the time the strongest pulse should be received in the correct room.

Equipment

- Two detectors in adjoining rooms
- P-Tags

Method

8 tags are placed in three separate rooms (rooms 1, 2, and 5; see Figure 4/Table 5) with varying orientation. The pulses are logged and the logs collected. They will then be analysed for the number of pulses detected and then tallied to work out the number of correct detections compared to the number of erroneous detections. For the purpose of analysis, all the zones in a room will be counted towards detection for the room. For instance in room #1, which has several zones the results, will be coded from the individual zones to rooms before analysis.

Result

The overall result is in strong support of hypothesis one. Only one tag out of eight had any detection that was erroneous in the data set. Upon closer examination the tag with the erroneous detections were oriented towards the outside wall (lined with windows). Due to the reconfigurable walls in the lab and the real walls that surround the reconfigurable cells, there is a minor gap (approximately 10-centimetres) between the wall and the windows which would not be there in a "real" room.

The tallies for the individual tags, as well as the combined total are shown in Table 7. All the other tags obtained 100% accurate (room-wise) detection throughout the test.

Tag ID	**Physical Position (Ref Figure 4)**	**Room #1** *(%)*	**Room #2** *(%)*	**Room #5** *(%)*	**N**	**Accuracy**
60460	#1	100%	-	-	406	100%
60461	#2		100%	-	466	100%
60462	#1	100%	-	-	419	100%
60463	#1	100%	-	-	364	100%
60464	#1	100%	-	-	473	100%
60465	#1	100%	-	-	471	100%
60466	#5		-	100%	484	100%
60467	#1	99,05% (n=418)	0,95% (n=4)	-	422	99,05%
				Total	*3505*	*99,9%*

Table 7 Tally of detected locations for 8 tags spread over three rooms

6.4.2. Zoned Room Precision and Accuracy

This test was designed to be conducted in the patient room #1, (also denoted #1 in Figure 4). To directly test the zone properties one detector was disconnected leaving the room divided into two equal halves.

The rationale for the first hypothesis is that tags should not "randomly" change location without actually physically moving. That is, a stationary tag should be detected consistently in the same location. The second hypothesis is a check of correctness.

Hypothesis

H1. The positioning of tags is stable over time.

H2. An individual tag is correctly placed >95% of detections, regardless of orientation.

Equipment

- 6 P-Tags, 10-second (+2 P-Tags for control group)

- 2 detectors

Method

After the disconnection of the third detector in room 1, the room will effectively be divided in two halves. In each of these halves 3 tags will be placed with varying orientation (facing left, right and directly towards the detector). This relatively low number of tags is intended to avoid any interference caused by bandwidth contention and thus starving or occluding tags from transmitting. The data should also be checked[25] for obvious starvation issues at the conclusion of the tests.

The three tags will be placed in a line at equal distance from the virtual mid-line dividing the room. The tag closest to the mid-line will be facing the mid-line (i.e. left/right depending on which half), followed by a tag facing the detector and finally a tag facing the closest wall (again left/right depending on which half). This enables capturing data for more than one orientation in a single experiment.

Two additional tags will be placed in adjacent rooms (rooms #2 and #5) to function as a control group verifying that the room-level precision and accuracy is maintained. It also gives a range in terms of strength to compare any potential wrongly placed detections.

Result

First attempting to perform the test with 8 tags (which is optimistical considering the results from the concurrency testing) some tags seemed to, as expected, be starved for extended periods of time (see the left graph in Figure 14). Removing one tag from each zone (i.e. two tags in total) yielded the plot on the right, while still some intermittent problems, no tags experienced prolonged starvation and hence the dataset is considered to be valid with respect to no starvation problems.

---

[25] By plotting a scatter graph of the detections per tag and looking for patterns of starvation.

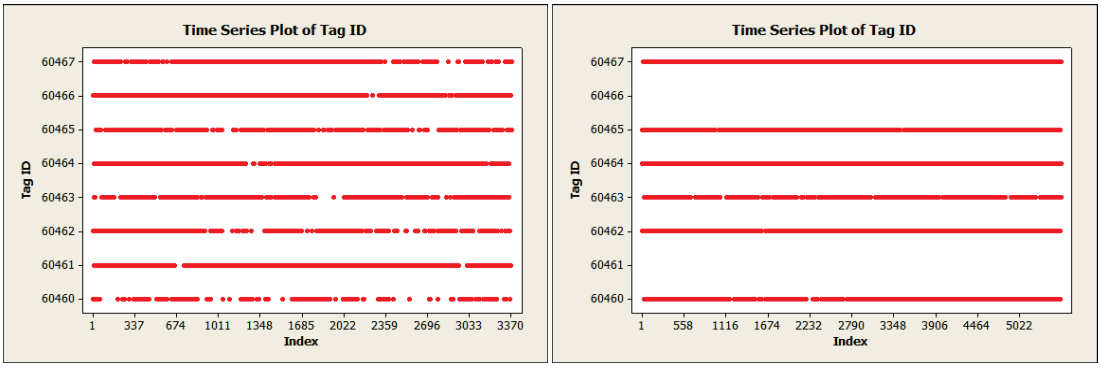


Figure 14 8-tag (left) and 6-tag (right) Zone Accuracy Time Series Plot

The three remaining tags for each zone were oriented so that one pointed directly towards the adjacent zone, one dead ahead towards the detector and the last tag were pointing towards the in-zone wall. This rationale behind this was to discover if reflections played an important part in the detection.

Tag Id	**Orientation**	**Accuracy**	**# Correct**	**# Erroneous**
60460	Out/Zone 1	**100%**	443	0
60462	Ahead/Zone 1	98%	451	7
60463	In/Zone 1	99%	405	3
60464	In/Zone 2	**78%**	400	113
60465	Ahead/Zone 2	95%	425	21
60467	Out/Zone 2	99%	460	3
Total		*95%*	*2584*	*147*

Table 8 Statistics for non-calibrated zone

Table 8 displays the results for evaluation of hypothesis 2, and we can see that the overall percentage of correctly placed tags is 95%, which was the lower bound of our hypothesis. However, it is noteworthy that excluding Tag 60464 (in Zone 2 oriented directly towards zone 1) brings this number up to 98%. On the whole a pretty good figure, but this also indicates strong evidence in support of refuting hypothesis 1 – the stability of the location over time. None of the tags in this test moved physically, but given the number of erroneous detections it is evident that the system interpreted changes in strength as movement. Figure 15 shows the location of Tag 60464 (in Zone 2, but with 78% accuracy). Each transition between 1 and 2 represents a move and we can also see that there are only a few periods (each pulse represents 10-seconds) in the test where the tag is detected as stationary during the whole test. This is an indication of lack of stability or precision in the detection.

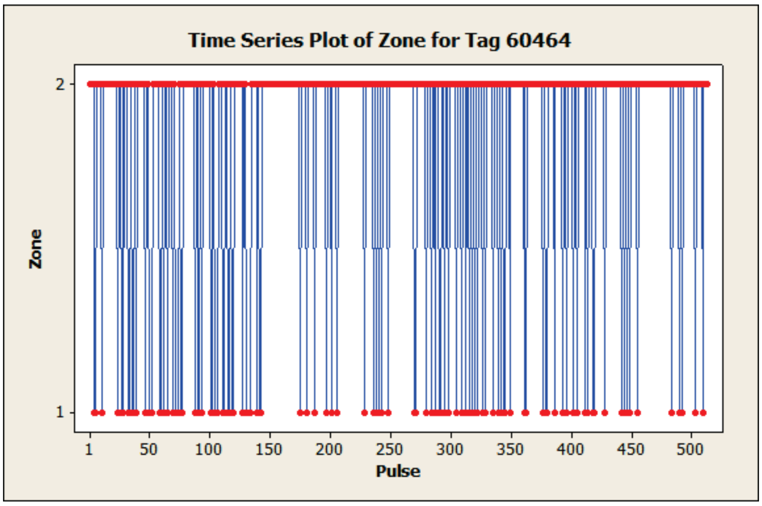


Figure 15 Zone for Tag 60464 Over Time

In a realistic real world installation, the detectors for adjacent zones would be calibrated, hence an extra round of experiments were in order. Using the first set of results, the average strength of detection was calculated for the tags on either side. This revealed (in line with the detector consistency testing) that there were a significant difference in the strength received by the two detectors (the averages were 21.7 and 15.8). Given this information a scaling factor of 0.72 (15.8 divided by 21.7) was applied to the stronger detector to obtain the second dataset (using the equipment and method, but with a different calibration).

Tag Id	**Orientation**	**Accuracy**	**# Correct**	**# Erroneous**
60460	Out/Zone 1	**99%**	403	3
60462	Ahead/Zone 1	92%	386	33
60463	In/Zone 1	**79%**	286	78
60464	In/Zone 2	94%	445	28
60465	Ahead/Zone 2	97%	457	14
60467	Out/Zone 2	95%	400	22
Overall	-	93%	2377	178

Table 9 Statistics for zone calibrated for equal average strength

With these two detectors having on average the same strength for their individual tags, the test was run again. The results are shown in Table 9. Interestingly enough, where as in the first test, the average accuracy for Zone 1 was approximately 99% it is now brought down to 90%. However, Zone 2 having 91% accuracy in test 1 now achieves 95%. This brings the overall accuracy down two percentage points to 93%. The tag with the lowest accuracy has also shifted from Zone 2 to the equivalent tag in Zone 1, hinting at the midline between the two detectors has shifted as well.

6.5. Evaluation of Test Results

For the first few baseline experiments there were bound to be variations in the results, expected from physical interference as well as minor variations in the components that make up the system.

Overall, and rather expectedly, the location system lives up to its basic promise, room-level accuracy. As found in the Room-level test (See Section 6.4.1), near 100% accuracy was obtained. This was even without any particular investment in calibration apart from carefully considered placement of the detectors to utilise the natural features of the rooms. The main caveat that realistically and most probably could be encountered in a production environment is the tag concurrency issues. This does not cause the system to fail completely, but can impact and degrade the performance of the system to such an extent that it might be wise to adopt some strategy for dealing with it (both in terms of software and/or design of the installation).

6.5.1. Dubious Strength Measurements

During both the consistency testing as well as sporadically in unrelated experiments, the system experienced what seemed to be random pulses that were detected at up to an order of magnitude stronger than both the immediately previous and subsequent pulses. One such measurement is captured in Figure 16, where the median for the whole series is 32.28, and the extreme value at index 10 measures 128.78, approximately four times stronger than both index 9 and 11. The tag remained stationary through the series; there were no other changing effects in the location system or in the room in the same period of time.

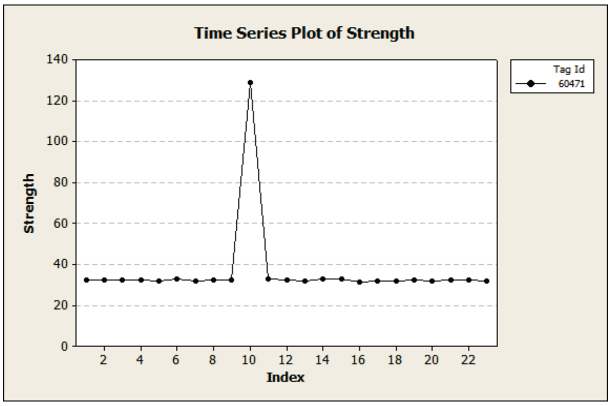


Figure 16 Time Series Plot of Strength with Spike

The study of tags and detectors uncovered no pattern for this rather quirky behaviour and given the available data does not support any good theory for their existence or reason.

A theory is that it is caused by glitches in the detector equipment or infrastructure. This is however not possible to confirm due to the “black box”-approach under which these experiments are performed. These glitches could range from external sources affecting cables to software/hardware issues in the DSPs processing the signal. The claim is backed up by observations of zoned areas where only one detector picks up this extreme variation in signal strength.

These spikes were not picked up by other adjacent detectors as one would expect (for instance when they occurred in overlapping/adjacent zones). If the tag did in fact transmit the pulse with such a change in strength, the change should have been detected by more than one detector where there were overlaps.

This behaviour is not a pronounced problem when occurring in room-level granularity installations, but in a zoned space this would cause the tag to potentially jump erroneously to adjacent zone (or even room) based on this extreme strength overruling what in reality was the correct measure. It was not observed that this caused the tag to be inaccurately located in the wrong room, but there is no evidence to suggest that it cannot happen in the case where there is sufficient strength that is picked up by near-by rooms (such as pulses transmitted while in overlaps from open doors or in doorways).

6.5.2. Consistency

The results from the consistency investigations were two-fold. Firstly the results are consistent to what is required for a functioning room-level system (and with some consideration, zone-level). The variations in the strength measurements limits the possibilities of using strength metrics for obtaining information about tag orientation or even more advanced scenarios such as using trilateration (possibly using angle-of-arrival estimations, based on strength) to obtain more fine grained location information.

Trilateration is possible, but given the restrictions imposed by the variations between different tags, and between successive measurements on individual tags, the value added by the detail in location is mitigated by the corresponding lack of precision and accuracy. In terms of capabilities this is something that most likely would be beyond what would be reasonable to use in a production-environment.

There is also circumstantial evidence for a minor processing of location data before feeding the location data to a middleware or application. The extreme values that occurred sporadically, as discussed in the preceding section, are detrimental to zoned locations and would have tags appear to be moving between adjacent zones whilst in reality the tag is stationary. This problem could be mitigating by defining a maximum value, either based on system performance (i.e. a pronounced change in signal strength over a short period of time) or by experiment reaching a practical maximum value for a normal performance and discarding measurements above this limit (similar to the function of a low-pass filter).

The detectors are in reality ultrasound microphones. As with any other microphones it is usually necessary to calibrate the microphones to create the same amount of open-circuit voltage given the same input. In this case the detectors were not explicitly calibrated by the experimenter, but used as they were provided. Whilst a full-scale implementation of a Sonitor IPS does not involve accurate (and scientific, see Pierce[43] for an excellent description for intra-microphone calibration without a pre-calibrated reference) calibration of microphones. The system allows for some rudimentary “calibration” where a system administrator can impose scaling factors on individual detectors to alter the balance between detectors.

The basic attempt at calibration (using an approximation Pierce’s method) as performed in Section 6.4.2 did not directly yield the desired effect in terms of increasing zoned accuracy. For an installation of a Sonitor IPS utilising zoned spaces, this would be an area to investigate further.

6.5.3. Range and Angle

Both hypotheses for this test were confirmed. This was rather unsurprising. The implication is that some simple ranging/lateration seems possible under strictly controlled circumstances. The variations between detectors and tags seem to suggest that this would purely be proof of concept rather than a reliable service due to the number of factors that needs to be controlled.

6.5.4. Concurrency

Concurrency is a parameter of a very restrictive nature, it defines who and what to tag as well as it has severe implications for rate of updates (hence the temporal resolution). The higher concurrency expected the lower rate of updates required. This scalability problem is not a unique issue with Sonitor, but an issue in all systems relying on centralised calculation of location (thus GPS and other LLC systems usually avoid this particular problem).

Using Inter-Arrival Time (IAT) as a Media Access Load Indicator

When tagging people it can be hard to avoid concurrency issues, in the sense that you cannot inform employees and patients not to gather in groups of more than a specific number of people in one space. Waiting areas can easily gather a lot of people. Hence it would be beneficial to obtain some way of detecting starvation of tags. Given that tags can transmit at different intervals as well as some tags fall asleep it can be hard to determine starvation automatically. A tag that has suddenly gone silent can just as likely be starved out of media access as it has left the tracked space.

Measuring IAT can to some extent predict whether or not tags are starved from transmitting, but it cannot reliably distinguish between the tag having left the area or being starved unless the system is in a steady state and non-starved state. Sonitor IPS does address the concurrency problem by the introduction of movement sensors in the E-series of tags, making the tags go into a sleep state when the tag has not moved for a certain number of transmissions (typically five or more) as well as having the tag transmit a “move”-bit making the infrastructure able to with a certain degree of accuracy distinguish between tags which has fallen asleep and one that has left the area.

The expected IAT from the tags in the system can either be manually entered into the information system when the tag is first introduced, or it is feasible to automatically deduce it based on the output metrics from the system (unless the system is operating at near capacity levels at all times, in case it would be difficult to obtain reliable data for calibrating the expected IAT for each tag).

Depending on the use of the location information the knowledge of expected IAT can either decrease the reliability (or accuracy, depending on use) of an earlier detection with time or register the tag as having unknown location (which in effect means changing the type of location, as per Dobson's taxonomy). Both approaches offer richer information than a more static system where a detection is taken as true until a new detection contradicts (or reaffirms) the current location. It is important to note here that it does not mitigate the underlying problem, it merely detects it.

The graphs in Figure 17 show effect on IAT measured over 2-minute windows at three different stages of the concurrency test (See section 6.3.4, with 2-minutes between introductions of new tags). The box ranges from first quartile through to the third quartile with the median indicated as a black line. The whisker extends to the upper limit[26] and lower[27] limit correspondingly. Values beyond the upper/lower limits (outliers) are indicated by stars. The tags used in this particular test were 10-second P tags, and the expected IAT would of course be 10-seconds (or 10.000-milliseconds as shown in the graphs).

For the first and second graph we can clearly see that the system is unsaturated, indicated by the fact that none of the IATs where larger than 11 seconds[28]. These two first graphs is in stark contrast to the last graph in Figure 17, where the upper limit for several of the tags extends beyond the 100-second marker and for several tags[29] the lower limit as well as the first quartile is larger than the expected value.

---

[26] Upper limit = $Q_3 + 1.5(Q_3 - Q_1)$

[27] Lower limit = $Q_1 - 1.5(Q_3 - Q_1)$

[28] The less than 11-second limit means that in reality there had been few-to-none collisions as a collision detection and media access negotiation would take IAT + transmission time which would be larger than 11 seconds.

[29] Notably tags 60463, 60468 and 60469

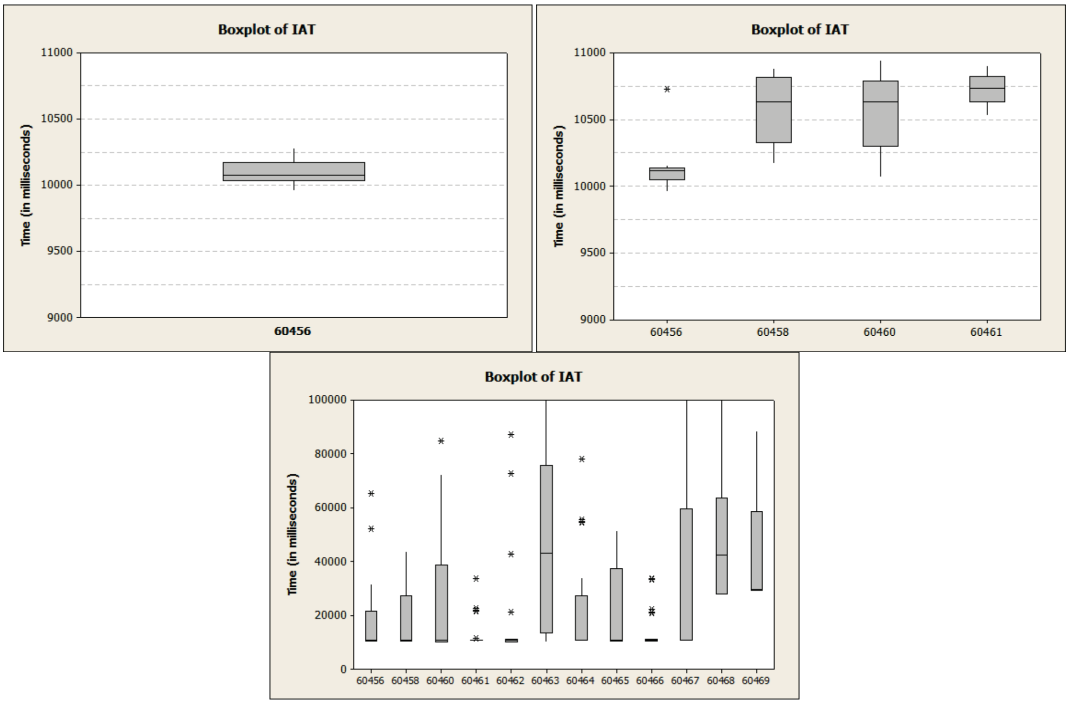


Figure 17 Box plots of IAT over 2-minute windows at T=0min, T=8min and T=12min, group by Tag ID

A proposed solution is to track the last n (for some appropriate measure of n) for all tags in the covered space and then detect any major deviations across several tags (to distinguish a tag that has left the space from starved tags, as starvation is likely to occur in all tags within the same space).

6.5.5. Detector Handover

For some use cases, a snapshot of location information is of secondary importance. The principal interest lies in tracking of movement over time and thus disconnected pieces of information is less valuable than a stream of changes in location for the tracked set of entities.

For most systems with piecewise or non-continuous infrastructure, the handover between detectors is highly important for the ability to track movement. Correspondingly in localized location computational systems the handover between the beacons providing clues has to function flawlessly to avoid disruptions in service when the client is moving from one beacon to another.

For indoor systems simple detector-handover occurs primarily in two situations. The simplest case is in the movement between two adjacent rooms, which both are covered by the location system. Especially with ultrasound, the fact that physical and virtual boundaries coincide means that the system is forced to deal with a detector handover as the detector left behind will often become out of range when entering a new room. More advanced, in terms of determining location, is the handover when both detectors remain in range, such as in the case of between zones or beams.

The Sonitor system operates with a temporal resolution in the order of several seconds. From the lowest level, where the tags take roughly 0.7-seconds to transmit their pulse, the system polls the base stations. Polling the base stations more frequently than the tags can transmit does not really serve any purpose. Additionally when taking into account that the tags in normal operation transmits at 5-second intervals or longer, means that the speed of which the handover can be detected is not really a meaningful metric for this system.

Trying to devise a test for this particular problem ended up being an exercise in design specification and reliable timing. The hard lower limit was the combination of polling rate of the server (which is configurable down to 1-second) combined with the pulse rate of the individual tags. Using this one can infer that the minimal temporal-resolution of the system is in the order of 1-second.

Validating this hypothesis against an external time source will inevitably have issues with timing of the experiment. However, when examining the inter-arrival times for the pulses on one tag (given non-saturated environment) one can see the typical variation in updates for one particular tag. For instance taking the first graph in Figure 17 one can see that the variation from the 10-second update interval is about 250-milliseconds from upper to lower extreme (a variation of about 3%). Given the low variation for a working system, the temporal resolution of the Sonitor system seems to be very close to the pulse rate of the tags rather than any features of the infrastructure.

7. Evaluation

In this chapter we will draw upon the work in Chapters 5 and 6 and see how the system testing performed in Chapter 6 can be used as valuable input for the capability assessment of a location system. The aim of the chapter is to evaluate the experimental approach to assessing an IPS, especially in terms of the identified capabilities from Chapter 5.

The initial sections draw upon Chapter 6 to form conclusions from the experiments and put them in the context of operational capabilities. The chapter will then use the synthesis of the theoretical and practical results to form an overall evaluation of the approach to assessing location systems.

7.1. From Theory to Capabilities

The literature search in Chapter 5, combined with knowledge of general location system characteristics from Chapter 3, helped devise the matrix formulated in 5.3.2. This matrix contains a description of *typical* purposes of location system implementations in hospitals. The dimension concept was then used to quantify the purposes into a grid, using the implications from context-aware systems theory as well as the description of the examples.

By including examples regardless of the technology used the aim is to achieve a tool that is technologically agnostic and makes up for the lack in precision by being very general and thus be useful for a larger set of questions. Similarly including all kinds of context-aware applications in the capabilities, makes the matrix usable to predict or extrapolate answers for purposes not directly covered in the matrix. The premise for being able to predict outside the range of the documented purposes lies with having a substantial understanding of existing solutions.

7.1.1. State of The Art

Looking at the matrix of the identified examples of location system use in the hospital, it is noticeable that the main number of systems lies in the lower 4 quadrants of the matrix. This might be an indication of the somewhat undeveloped state of location system use in the domain. There is a lack of systems in the higher quadrants of either dimension. The identified papers are all rather basic applications limited to one or a few purposes rather than using the gathered context for more advanced context-aware systems.

While there is a lot of interest in location systems from within the hospitals too, this is perhaps an indication that more advanced applications have yet to arrive.

7.2. From Experiments to Capabilities

Approaching the operational capabilities from the other side, assessing the capabilities of a specific location system rather than to define the location system based on them, we used the same location system knowledge to devise experiments to extract comparable information about the system.

To keep the assessment matrix free of technology specific dimensions, the idea behind the experiment design was to keep them close to the purposes, while at the same time keeping them simple and small scale.

Identifying how the dimensions map onto the characteristics of the Sonitor IPS, several tests were designed and performed as described in Chapter 5. A summary of these experiments is shown in Table 10 with the prediction of how they measure in terms of the dimensions.

Dimension	Test	Result
Granularity	- Room-Level Precision and Accuracy (see 6.4.1) - Zoned Room Precision and Accuracy (see 6.4.2) - Range and Angle of Arrival (see 6.3.3)	Low to medium-high.
Resolution	n/a Found theoretically (see 6.5.5)	Low to Medium.
Concurrency	Tag Concurrency (see 6.3.4)	Expression for calculating maximum concurrency based on number of tags (see 6.3.4)

Table 10 Results from experiments

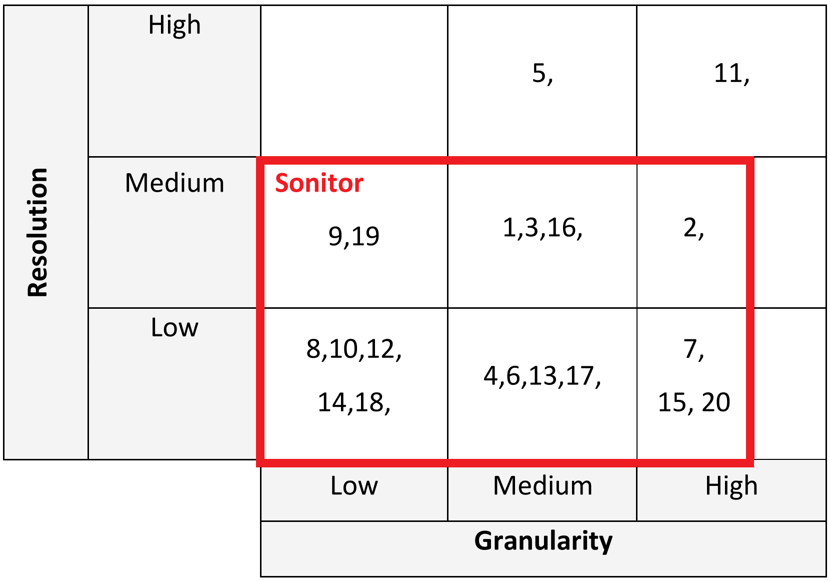
Of the two dimensions that were used for classification, resolution was the simpler find for the Sonitor equipment. Sonitor leaves few options in how one can influence the resolution directly. The hard lower limit is the time for one pulse to transmit and as such it places itself rather in the middle of the resolution dimension with a resolution of minimum in the order of seconds (see table in Chapter 5.2.2). With P-Tags that transmit even less frequently, the Sonitor system has lower-medium to low resolution.

In terms of granularity Sonitor Technologies makes rather bold statements claiming it “tracks real-time location of moveable equipment or people in complex indoor environments with 100% room-level, or zone-level accuracy (such as bed-level) within a room”[37]. The experiments in Chapter 6 confirmed their room-level claim reaching similar figures without any problems. In 6.4.2, where the zone-level performance was tested directly, the measured accuracy was 95%, which is to say less than completely reliable (the rudimentary attempt at improving the accuracy actually had the opposite effect which lends some weight towards accepting that the first figure is in the upper bounds of what is possible).

Extrapolating these experiments to determine accuracy onto granularity we propose that the Sonitor equipment stretches into high-granularity territory according to the determined scale, but does not possess all the qualities to include the whole high-granularity area.

7.2.1. Applying results to the Matrix

Applying these values to the Matrix devised in Chapter 5, yields Table 11.



Resolution	High		5,	11,
	Medium	**Sonitor** 9,19	1,3,16,	2,
	Low	8,10,12, 14,18,	4,6,13,17,	7, 15, 20
		Low	Medium	High
		Granularity		

Table 11 Matrix with Superimposed Test Results

Here we can see that the Sonitor equipment posses most of the capabilities and as such would be a technically good fit for challenges in the lower 4 quadrants of the matrix. The red rectangle also points out that the Sonitor equipment does not possess the capabilities to solve points 5 and 11.

According to the dimensions examined Sonitor does not possess the capability to perform “Locating the nearest staff to a particular room”(5) at the required resolution for it to be useful in a context-aware system. The “Auto Log on/off hospital computer systems”(11) purpose, is for the Sonitor system outside the range of both dimensions. Certain tags lacked precision in the detected location (See the results in 6.4.2) when positioned in zones. That would in the case of automatic log off mean that the user would be logged off without actually having left the computer. Along the granularity dimension, the feasible smallest zone size of the Sonitor system would make it difficult to have several computers in one room.

7.2.2. Granularity and Size of Infrastructure

As briefly discussed in Chapter 5, granularity and size of infrastructure are dependent on each other, and the use in the capability assessment was rather indistinct about which factor is causing the effect. However as a first triage of systems, and to keep the assessment tool flexible and technology-agnostic, it seems necessary to keep these factors together.

In later stages of an assessment when the focus is on one or two systems and the questions focus less on which system and more on the required size of the installation it might be wise to attempt to decouple the two. One way to decouple them is to repeat the experiments (and find potentially new) values for the dimensions with using the predicted size of infrastructure to perform the experiments. If the experiments do not support the required granularity, one would go back and change the prediction and repeat.

If there is a practical upper (and lower) density of infrastructure, those should be tested first to establish the span of granularity to confirm whether or not the required level is obtainable.

7.2.3. Required Number of Tests

The number of tests performed when attempting to experimentally assess the location system is larger than the number used for instances in fixing the results in section 7.2.1. That is not to say that any additional tests were unnecessary, rather to the contrary. A good understanding of the location system improves the accuracy of the prediction in terms of dimensions. Further it is relatively cheap both time-wise and economically to perform experiments of limited size in the comfort of the laboratory rather than going ahead with a pilot and discovering any problems.

The problem with the rather coarse assessment tool as it stands now is the lack of support for classifying the knowledge gained from these tests that does not directly map onto any dimension of the matrix.

7.3. Value of Such an Assessment Tool

One of the main strengths of the operational capabilities is their “plain language” form. This enables people without in depth expertise in location systems to participate in the design choices in a more informed manner. For a domain such as healthcare where a lot of the demands for new software or technology comes from actors without engineering backgrounds, such a tool could be both a facilitator for development of new ideas as well as a guide in realising the full potential of existing possibilities.

Further it is flexible in the sense that it uses a loose definition of location systems as the foundation for the theory, hence not excluding systems that for instance would not be considered using a stricter (for instance ISO/IEC) definition of location systems. The focus is on purpose and merits rather than technological methods.

7.4. Cost

While cost is not a technical factor, it is perhaps the most important non-technical factor when assessing a location system and should not be ignored completely even in a technical evaluation. Cost is also tightly connected to the various systems and is difficult to compare across different systems.

Fixing dimensions for the various purposes at the lowest possible place was partly driven by attempting to permit cost comparisons. In the case of the granularity dimension it would be wise to attempt to calculate cost-curves for the various levels and see where the return on investment levels out where it will become economically unsound to increase the infrastructural density further.

7.5. How to Utilise the Method in Practice

Combining sections 7.1 and 7.2 we can obtain a method that spans from theory (purpose) via operational capabilities to experiments (location system). This is perhaps the most practical useful result of this work. For the hospital wishing to utilise the results of this Thesis there are two main ways to utilise this assessment tool.

If the design process has a particular purpose it wants to tackle with a location system, the challenge is to identify which location system is appropriate to deploy. Faced with this challenge the procedure is to identify where this purpose lies within the matrix and then seek out a location system that has the operational capabilities to match it.

Conversely, if the location system has been identified, the matrix method can be combined with experiments to obtain which quadrants the location system supports. Through this one can identify applications that can be realised using the given location system.

Using these simple dimensions it would seem beneficial for hospitals to share their findings with each other. This would increase the usability of the method for each application of it, as well as reduce the amount of work duplicated across hospitals. After pilot implementations, the hospital should aim at publishing the results from the implementation to enable other institutions to take advantage of the work and results.

7.5.1. From Purpose to Location System

With a good description of the purpose it will either be possible to identify similar challenges in the existing matrix and adjust the dimensional values accordingly to fit the new purpose to find in which quadrant of the matrix this purpose belongs. If the purpose does not match any other examples, it will be necessary to examine it closer to find appropriate values for the dimensions. This involves examining the purpose more closely to understand how it is affected by the dimensions used in the assessment.

When this is complete, the purpose can be placed in the matrix. The next aim is to find a location system that supports the capabilities required. An obvious first step here would be to look at the systems already in place and investigate whether any of these have any location system qualities and if so attempt to experimentally obtain values to compare to the dimensions of the sought purpose.

If the purpose cannot be solved within existing infrastructure, the search continues with new systems. Using the matrix it should be possible to narrow the focus rather than searching through all location systems by focusing on systems that either has shown capabilities in the same quadrant in literature or through promises from the prospective vendors. The last step would be to experimentally re-test the capability of the location system to ensure that it operates as expected in the target environment.

7.5.2. From Location System to Opportunities

In the opposite situation where the location system is already identified, either because one wants to expand the number of tasks it supports or because one is attempting to improve the utilisation of existing infrastructure for locative purposes, the assessment tool can be used.

The first step is to perform experiments on the location system to find the appropriate dimensional values as to see where it fits in the existing matrix. In the somewhat unlikely situation that the hospital does not have any particular purposes it wishes to fulfil, the matrix can offer suggestions that are within reach of the installed location system.

More likely it can offer suggestions to whether or not the location system has to be extended or improved in order to align the operational capability of the system and the capability requirements of the sought purpose. In this case possibly even identifying which properties have to be improved (e.g. granularity, concurrency, or resolution).

7.6. Summary

The operational capabilities matrix devised bridges the gap between purpose of a system, capabilities required and the collection of the same capabilities. We have shown how information gathered experimentally on a Sonitor IPS can be used to extrapolate the capabilities of the system which in turn can describe the suitability of a Sonitor IPS installation for fulfilling various purposes in the hospital domain.

The chapter also summarises and shows how this method can be applied by a hospital either seeking to identify a location system based on a particular need or how to identify opportunities based on an existing location system.

8. Conclusion

This thesis has shown how to extract and compile a matrix of operational capabilities for a location system in a particular domain. In terms of what operational capabilities need to be understood before designing an indoor location system implementation (RQ1); the dimensions to be found and thoroughly understood are: granularity, resolution and concurrency.

To assess these operational capabilities experimentally (RQ2) a small-scale pilot installation of a location system was used to determine how this particular location system compares to the identified dimensions through simple low cost/effort experiments. It was also discussed and shown how categorising existing examples of implementations using the same dimensions, could extrapolate the feasibility of the location system on to similar purposes.

Using the same method used for assessing the capabilities experimentally we found that it is possible to obtain predictions of the impact of infrastructure size on the overall operational capabilities (RQ3) by repeating the experiments for either end of the possible infrastructural density to find a range. With this range it is possible to create a reasonably reliable upper and lower granularity prediction.

The exploration also reiterated that the use of location technology in healthcare is in its infancy. The majority of the identified examples from the domain came in reasonably low requirements in terms of the investigated dimensions, making them rather basic applications. The approach used in this Thesis was highly domain specific. It is however likely that this approach can be just as valuable for other domains with similar aspects as the healthcare domain.

9. Further Work

There are several areas in which there is possible further work.

The most obvious work is to extend the matrix by introducing more dimensions and compare whether the additional information does in fact improve the precision of the predictions. Similarly one could attempt to vary the amount of examples put into the matrix and evaluate the effect of increasing it.

The assumption is that larger amounts of examples will make the matrix easier to use. This might also improve the imbalance between the various quadrants. As of now there are few examples in the highest quadrants, does this validate the claim that location systems in healthcare is in its infancy, or is it due to inaccuracies in the literature search methodology?

The most interesting proposition, to the author, would be to attempt to assess another location system using the same matrix. This would provide an opportunity to validate the method by providing results that can be compared against the Sonitor experiments.

This could also provide input into the location system knowledge base that was described in Chapter 7.

The final point would be to show that the method does indeed transfer to other domains by apply in different a different setting, using different literature as theoretical input.

References:

1. Harrop, P., G. Holland, and R. Das, *Real Time Locating Systems 2009-2019*. 2009, IDTechEx.
2. Lin, C.C., et al., *Wireless health care service system for elderly with dementia.* IEEE Trans Inf Technol Biomed, 2006. **10**(4): p. 696-704.
3. Marjamaa, R., et al., *Time accuracy of a radio frequency identification patient tracking system for recording operating room timestamps*. 2006, IARS. p. 1183-1186.
4. Rogers, A., E. Jones, and D. Oleynikov, *Radio frequency identification (RFID) applied to surgical sponges.* Surg Endosc, 2007. **21**(7): p. 1235-7.
5. Sandberg, W.S., et al., *Automatic detection and notification of "wrong patient-wrong location" errors in the operating room.* Surg Innov, 2005. **12**(3): p. 253-60.
6. Hagland, M., *Follow that pump. As medical devices grow more expensive, tracking them becomes absolutely imperative.* Healthc Inform, 2009. **26**(3): p. 42, 80.
7. Meyer, M., et al., *Automatic Time-Motion Study of a Multistep Preoperative Process. Economics.* Anesthesiology, 2008. **108**(6): p. 1109-1116.
8. Faxvaag, A., A.R. Seim, and P. Toussaint, *COSTT Presentation given at HelsIT 2008*. 2008.
9. Dey, A. and G. Abowd, *Towards a better understanding of context and context-awareness.* 1999.
10. Winograd, T., *Architectures for context.* Human-Computer Interaction, 2001. **16**(2): p. 401-419.
11. Brown, P., J. Bovey, and X. Chen, *Context-Aware Applications: From the Laboratory to the Marketplace.* IEEE Personal Communications, 1997. **4**(5): p. 58-64.
12. Dey, A.K., *The CyberDesk Project.* AAAI 1998 Spring Symposium on Intelligent Environments, Technical Report SS-98-02, 1998. **11**: p. 3-13.
13. Ryan, N., J. Pascoe, and D. Morse, eds. *Enhanced Reality Fieldwork: the Context-Aware Archaeological Assistant.* Computer Applications in Archaeology, ed. V. Gaffney, M. van Leusen, and S. Exxon. 1997.
14. Schilit, B. and M. Theimer, *Disseminating Active Map Information to Mobile Hosts.* IEEE Network, 1994. **8**(5): p. 22-32.
15. ISO/IEC, *Automatic identification and data capture (AIDC) techniques - Harmonized vocabulary - Locating systems*, in *Information Technology*. 19762-5:2008(E), ISO: Switzerland.
16. Dobson, S., *Leveraging the subtleties of location*, in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*. 2005, ACM: Grenoble, France.
17. Indulska, J. and P. Sutton. *Location management in pervasive systems*. 2003: Australian Computer Society, Inc. Darlinghurst, Australia, Australia.

18. Hightower, J. and G. Borriello, *A Survey and Taxonomy of Location Systems for Ubiquitous Computing*. 2003, University of Washington (UW), Computer Science and Engineering: Seattle.
19. Hightower, J. and G. Borriello, *Location sensing techniques.* IEEE Computer, 2001.
20. Jones, R.P. and D.H. Conway, *The effect of electromagnetic interference from mobile communication on the performance of intensive care ventilators.* Eur J Anaesthesiol, 2005. **22**(8): p. 578-83.
21. Wallin, M.K., T. Marve, and P.K. Hakansson, *Modern wireless telecommunication technologies and their electromagnetic compatibility with life-supporting equipment.* Anesth Analg, 2005. **101**(5): p. 1393-400.
22. Hohlrieder, M., et al., *The impact of avalanche transceivers on mortality from avalanche accidents.* High Altitude Medicine & Biology, 2005. **6**(1): p. 72-77.
23. Brightkite. *Accessed 2009-03-04*. Available from: http://www.brightkite.com.
24. Government Accountability Office, U.S., *Global Positioning System - Significant Challenges in Sustaining and Upgrading Widely Used Capabilities (GAO-09-325)*, U.S.G.A. Office, Editor. 2009.
25. Maltz, J., et al. *The Trauma Patient Tracking System: implementing a wireless monitoring infrastructure for emergency response*. 2005.
26. Konomi, S. and G. Roussos, *Ubiquitous computing in the real world: lessons learnt from large scale RFID deployments.* Personal and Ubiquitous Computing, 2007. **11**(7): p. 507-521.
27. Prince, P., *United states sets date for e-passports.* The RFID Journal, 2005. **25**.
28. Hightower, J., G. Borriello, and R. Want, *SpotON: An indoor 3D location sensing technology based on RF signal strength*, in *UW CSE 00-02-02, University of Washington, Department of Computer Science and Engineering, Seattle, WA*. 2000.
29. Ward, A., A. Jones, and A. Hopper, *A new location technique for the active office.* Personal Communications, IEEE, 1997. **4**(5): p. 42-47.
30. Minami, M., et al., *DOLPHIN: a practical approach for implementing a fully distributed indoor ultrasonic positioning system.* Lecture notes in computer science, 2004: p. 347-365.
31. Harle, R. and A. Hopper. *Deploying and evaluating a location-aware system*. 2005: ACM New York, NY, USA.
32. Holm, S., et al. *Indoors data communications using airborne ultrasound*. in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*. 2005.
33. Priyantha, N., A. Chakraborty, and H. Balakrishnan. *The cricket location-support system*. in *6th ACM International Conference on Mobile Computing and Networking (ACM MOBICOM)*. 2000. Boston, MA: ACM New York, NY, USA.
34. Priyantha, N.B., *The Cricket Indoor Location System*, in *Department of Electrical Engineering and Computer Science*. 2005, Massachusetts Institute of Technology: Boston.
35. Krohn, R., *The Optimal RTLS Solution for Hospitals - Breaking Through a Complex Environment.* Journal of Healthcare Information Management, 2008. **22**(4).
36. Grey, M.M., *Ultrasound's unique use. Seton Medical Center is using ultrasound to do more than image the heart; it's tracking clinical assets and patients.* Healthcare informatics, 2007. **24**(12): p. 29-32, 34.

37. Sonitor, *RTLS in Hospitals (presentation)*, in http://www.sonitor.com, Sonitor, Editor. 2008.
38. Laskowski, M., et al., *Models of emergency departments for reducing patient waiting times.* PLoS One, 2009. **4**(7): p. e6127.
39. Miodownik, S., P.H. Frisch, and P. Booth, *Application of RFID in an integrated healthcare environment.* IEEE Engineering in medicine and biology society conference proceedings, 2006. **1**: p. 117-9.
40. Sandberg, W.S.W., et al., *Using location tracking data to assess efficiency in established clinical workflows.* AMIA ... Annual Symposium proceedings, 2006: p. 1031.
41. Wicks, A.M.A., J.K.J. Visich, and S.S. Li, *Radio frequency identification applications in hospital environments.* Hospital Topics, 2006. **84**(3): p. 3-8.
42. Miller, M., *Tuning into future healthcare use of RFID.* Automatic I.D. News, 1999. **15**(2): p. 58.
43. Pierce, A.D., *Acoustics : an introduction to its physical principles and applications*. 1989 ed. 1989, Woodbury, N.Y.: Acoustical Society of America. xxv, 678 p.

A. Descriptive Statistics for Tag Consistency

30-Second P-Tags:

Tag Id	N	Average	Median	Max	Min	StdDev
60470	102	13,719	13,968	17,223	9,184	1,107
60471	93	7,892	8,038	8,666	5,863	0,538
60472	99	10,532	10,629	11,779	7,571	0,685
60473	106	12,113	12,241	13,728	8,500	0,828
60474	99	11,192	11,322	12,439	8,198	0,778
60475	100	9,302	9,441	10,609	5,931	0,753
60477	111	8,473	8,566	9,155	6,040	0,455
60478	104	10,120	10,274	13,613	6,742	0,818
Sum	*814*	*10,42*				

10-Second P-Tags:

Tag Id	N	Average	Median	Max	Min	StdDev
60461	587	8,004	7,963	9,179	7,368	0,275
60462	393	9,673	9,635	11,437	8,951	0,360
60463	384	9,697	9,679	11,252	8,763	0,378
60464	530	8,547	8,523	9,969	7,665	0,344
60465	401	7,375	7,343	8,488	6,691	0,237
Sum	*2295*	*8,659*				

7/10-Second E-Tags:

Tag Id	**N**	**Average**	**StdDev**	**Median**	**Max**	**Min**
6372	67	17,840	2,313	17,767	23,986	13,751
6375	67	19,309	1,573	19,312	22,031	15,290
6376	67	20,226	2,153	20,004	24,833	14,613
6377	69	17,388	1,818	16,917	20,950	13,941
6380	80	18,202	1,528	18,195	21,839	14,702
6381	71	20,711	2,203	20,466	25,851	16,027
6382	81	18,808	1,987	18,657	25,384	14,706
6386	89	21,865	1,125	21,837	24,403	18,482
6387	86	21,297	2,325	21,310	26,288	15,953
6391	89	23,758	2,496	23,845	37,153	17,250
Sum	*766*	*19,940*				

B. Descriptive Statistics for Range and Angulation

	A of I	-135	-90	-45	0	45	90	135	180
50 cm	Mean	3,324846	14,11393	40,90327	58,47146	32,27978	19,11455	3,616113	2,465477
	Max	3,647325	14,96848	44,14209	59,99088	32,73373	19,4993	3,836317	2,744275
	Min	3,021676	13,37053	39,58074	57,30568	32,0051	18,55594	3,373317	2,293249
	Std.Dev	0,17608	0,637277	1,610289	0,751691	0,248161	0,295398	0,124095	0,135139
	Num	8	10	10	10	8	9	10	9
100cm	Mean	1,671816	6,691273	19,28051	26,9662	18,22006	8,843207	3,068595	1,847034
	Max	1,803905	6,91507	19,62204	27,51539	18,82655	9,197143	3,166563	2,008264
	Min	1,536682	6,495776	18,93173	26,41058	17,89736	8,650723	2,956217	1,67179
	Std.Dev	0,091821	0,139267	0,24695	0,365213	0,255673	0,189095	0,059368	0,117764
	Num	8	10	10	10	10	10	10	9
150cm	Mean	1,300916	2,292371	13,0615	16,6561	12,94337	3,510022	2,138006	1,564508
	Max	1,443284	2,555084	13,45076	17,27171	13,75291	3,64034	2,307618	1,717691
	Min	1,165284	2,060551	12,57086	16,17947	12,38786	3,359347	1,976333	1,471822
	Std.Dev	0,116088	0,179193	0,288704	0,38288	0,385034	0,099056	0,106989	0,078305
	Num	8	10	10	10	10	10	9	7
200cm	Mean	1,445793	3,668979	6,543453	9,72179	6,440714	2,766687	1,715753	1,256003
	Max	1,580388	3,875633	6,81329	10,16964	6,679578	3,185323	1,781154	1,41255
	Min	1,301789	3,482481	6,435106	9,013539	6,222764	2,484835	1,648441	1,138741
	Std.Dev	0,103989	0,1199	0,115383	0,385341	0,138344	0,244349	0,059661	0,090077
	Num	7	10	10	10	10	10	7	7

C. Purpose-made scripts

For reformatting to a more Excel friendly format:

```perl
#!/usr/bin/perl

# Written by Andreas D. Landmark <andreala@idi.ntnu.no>
# For use in parsing logs from SonitorIPS Server into more excel friendly format
#

use warnings;


while (<>) {
        my $string = $_;
        my @split_str;
        @split_str = split(",", $string);

        # Time
        print $split_str[3] . ":" . $split_str[4] . ":" . $split_str[5] . "\t";

        # milliseconds for precise timing: "." . $split_str[6] . "\t";

        # Date
        print $split_str[2] . "-" . $split_str[1] . "-" . $split_str[0] . "\t";

        # Detector
        print $split_str[8] . "\t";

        # Strength
        $strength = $split_str[11] . "\t";
        # Replacing . with , to comply with Norwegian decimal number formatting
        $strength =~ s/\./,/g;
        print $strength;

        # Confidence
        $confidence =  $split_str[12] . "\t";
        # Replacing . with , to comply with Norwegian decimal number formatting
        $confidence =~ s/\./,/g;
        print $confidence;

        # Selected
        print $split_str[19];

```

For calculating Inter-Arrival Times(IAT)

```
package no.costt.SonitorIPS.LogParsers;

import java.io.BufferedReader;
import java.io.FileNotFoundException;
import java.io.FileReader;
import java.io.IOException;
import java.util.ArrayList;
import java.util.Calendar;
import java.util.TreeMap;


/**
 * Class written to parse default Detector logs from Sonitor IPS
 * server for IAT Analysis
 *
 * @author Andreas Landmark <andreala@idi.ntnu.no>
 *
 */
public class IATAnalysis {
       public static void main(String[] args) {
              System.out.println("Sonitor IPS Detector Log Parser - IAT Analysis");
              System.out.println("Andreas D. Landmark <andreala@idi.ntnu.no> April
2009");

              if(args.length == 0) {
                     System.out.println("Usage:\n DetectorLogsParser <filename> [>
output.file]\n\n");
                     System.out.println("The parser outputs statistics on stderr, so
be sure to redirect either stdout or stderr");
                     return;
              }


              FileReader fr;
              int numOfDet = 0;

              try {
                     fr = new FileReader(args[0]);
                     BufferedReader br = new BufferedReader(fr);
                     String line;

                     TreeMap<String, TreeMap<String, ArrayList<Calendar>>> detectors
= new TreeMap<String, TreeMap<String, ArrayList<Calendar>>>();

                     line = br.readLine();

                     do {
                            numOfDet++;
                            //                                 Sample input
                            //
       2009,4,21,8,38,28,501,60475,"Detector
1",129.241.172.56,4,8.053412,11355.000000,1,0,0,0,0,0,1
                            //
       2009,4,21,8,38,28,511,60475,"Detector
2",129.241.172.56,5,7.716141,13956.000000,1,0,0,0,0,0,0

                            String[] parts = line.split(",");


                            // Using fields 9+10 as key id, gives IP(9):portNo(10)
                            String detId = parts[9] + ":" + parts[10];

```

```
                if(detectors.containsKey(detId)) {
                    if(detectors.get(detId).containsKey(parts[7])) {

    detectors.get(detId).get(parts[7]).add(IATAnalysis.calFormatter(parts));
                    } else {
                        ArrayList<Calendar> cals = new
ArrayList<Calendar>();
                        cals.add(IATAnalysis.calFormatter(parts));
                        detectors.get(detId).put(parts[7], cals);
                    }
                } else {
                    ArrayList<Calendar> cals = new
ArrayList<Calendar>();
                    cals.add(IATAnalysis.calFormatter(parts));
                    TreeMap<String, ArrayList<Calendar>> tags = new
TreeMap<String, ArrayList<Calendar>>();
                    tags.put(parts[7], cals);
                    detectors.put(detId, tags);
                }

                line = br.readLine();
            } while(line != null);
            System.err.println("Found: " + detectors.size() + "
detector(s)");
            for(String detId : detectors.keySet()) {
                System.err.println("For Detector \"" + detId + "\" found
" + detectors.get(detId).size() + " tags");
                for(String tagId : detectors.get(detId).keySet()) {
                    System.err.println(" |-tag" + tagId + "-> Found " +
detectors.get(detId).get(tagId).size() + " detections ");
                }
            }

            // Now we collate and print...

            ArrayList<ArrayList<String>>x = new
ArrayList<ArrayList<String>>();
            for(String detId : detectors.keySet()) {
                for(String tagId : detectors.get(detId).keySet()) {
                    ArrayList<String> y = new ArrayList<String>();
                    y.add(detId);
                    y.add(tagId);
                    Calendar prevCal = null;
                    for(Calendar myCal :
detectors.get(detId).get(tagId)) {
                        if(prevCal != null) {

    y.add(Long.toString(myCal.getTimeInMillis() - prevCal.getTimeInMillis()));
                        }
                        prevCal = myCal;
                    }
                    x.add(y);
                }
            }

            int errCount = 0;
            String outLine = "";
            for(int j = 0; j < 10000; j++) {
                errCount = 0;
                for(int i = 0; i < x.size(); i++) {
                    try {
                        outLine = outLine.concat(x.get(i).get(j) +
";");
                    } catch (IndexOutOfBoundsException e) {
                        outLine = outLine.concat(";");
                        errCount++;
                    }
                }
```

```
				if(errCount >= x.size()) {
					return;
				}
				System.out.println(outLine);
				outLine = "";
			}



		} catch (FileNotFoundException e) {
			System.out.println("ERROR: Could not find the file " + args[0] +
". Exiting...");
			return;
		} catch (IOException e) {
			// TODO Auto-generated catch block
			//				e.printStackTrace();
		}

	}


	/**
	 * Takes in a log of default formatted SonitorIPS Server Log file, and
returns a java.util.Calendar object populated with the same date
	 * @param parts a unformatted log line
	 * @return Calendar object populated with the information from the log line
	 */
	public static Calendar calFormatter(String[] parts) {
		Calendar myCal = Calendar.getInstance();
		myCal.set(Integer.parseInt(parts[0]), Integer.parseInt(parts[1]),
Integer.parseInt(parts[2]), Integer.parseInt(parts[3]), Integer.parseInt(parts[4]),
Integer.parseInt(parts[5]));
		myCal.set(Calendar.MILLISECOND, Integer.parseInt(parts[6]));

		return myCal;
	}
}
```

```
package no.costt.SonitorIPS.LogParsers;

import java.io.BufferedReader;
import java.io.FileNotFoundException;
import java.io.FileReader;
import java.io.IOException;
import java.text.NumberFormat;
import java.util.Locale;


/**
 * Class written to parse default Detector logs from Sonitor IPS
 * server for Detector Consistency tests.
 *
 * @author Andreas Landmark <andreala@idi.ntnu.no>
 *
 */
public class DetectorParser {


	public static void main(String[] args) {
		System.out.println("Sonitor IPS Detector Log Parser");
		System.out.println("Andreas D. Landmark <andreala@idi.ntnu.no> April
2009");

		if(args.length == 0) {
			System.out.println("Usage:\n DetectorLogsParser
<filename>\n\n");
			return;
		}

		try {
			FileReader fr = new FileReader(args[0]);
			BufferedReader br = new BufferedReader(fr);
			String line;




			int prevTime = 0;
			int prevTag = 0;
			int prevDetector = 0;
			double prevStrength = 0.0;
			int numOfDet = 0;
			int numOfDDet = 0;

			int rightDetector = -1;

			boolean header = true;
			line = br.readLine();

			NumberFormat numberFormatter =
NumberFormat.getNumberInstance(Locale.getDefault());

			do {
				numOfDet++;
				try {
					//				Sample Input
					//
	2009,4,21,8,38,28,501,60475,"Detector
1",129.241.172.56,4,8.053412,11355.000000,1,0,0,0,0,0,1
					//
	2009,4,21,8,38,28,511,60475,"Detector
2",129.241.172.56,5,7.716141,13956.000000,1,0,0,0,0,0,0

					String[] parts = line.split(",");

					int tag = Integer.parseInt(parts[7]);
```

```
                    int time = Integer.parseInt(parts[5])*1000 +
Integer.parseInt(parts[6]);
                    int detector = Integer.parseInt(parts[10]);
                    double strength = Double.parseDouble(parts[11]);


                    if(prevTime == 0 || prevTag == 0) {
                        prevTime = time;
                        prevTag = tag;
                        prevDetector = detector;
                        rightDetector = detector; // Set to ensure
consistent sorting of the columns
                        prevStrength = strength;
                    } else {
                        if(tag == prevTag && detector !=
prevDetector) {
                            // Check if time is near enough to be
double det
                            // Time doesn't seem to change the
number of det, even up to values of 30s
                            if(time - prevTime < 1500) {
                                numOfDDet++;
                                //
                System.out.println("Double detection:");
                                if(header) {

    System.out.println("hours;minutes;seconds;milliseconds;tagId;detectorA;streng
thA;detectorB;strengthB");
                                    header = false;
                                }

                                // sorting them in the same
order... at the same time, mixing up the time stamps
                                // however the difference in
them should be negligible
                                if(detector == rightDetector) {

    System.out.println(parts[3] + ";" + parts[4] + ";" + parts[5]+ ";" +
parts[6]+ ";"
                                            + tag + ";"
+ detector + ";" + numberFormatter.format(strength) + ";" + prevDetector + ";"
                                            +
numberFormatter.format(prevStrength));
                                } else {

    System.out.println(parts[3] + ";" + parts[4] + ";" + parts[5]+ ";" +
parts[6]+ ";"
                                            + tag + ";"
+ prevDetector + ";" + numberFormatter.format(prevStrength) + ";" + detector + ";"
                                            +
numberFormatter.format(strength));

                                }
                            }
                        }
                        prevTime = time;
                        prevTag = tag;
                        prevDetector = detector;
                        prevStrength = strength;


                    }

                    line = br.readLine();
                } catch (IOException e) {
                    //
    e.printStackTrace();
                    line = null;
                }
```

```
                } while(line != null);

                System.out.println("Number of detections in total: " +
numOfDet);
                System.out.println("Number of double detections: " + numOfDDet);

            } catch (FileNotFoundException e) {
                System.out.println("Couldn't find the file: " + args[0]);
            } catch (IOException e) {
                System.out.println("Empty file, exiting...");
            }
        }
```