# Device-free, Radio-based Activity Recognition using Smart Home Wireless Communication Technologies

To Attain the Academic Degree of

Doctor of Engineering

Faculty of Informatics
Karlsruhe Institute of Technology (KIT)

**approved**

**Dissertation**

by

**Markus Scholz**

Born in Leipzig, Germany

Examination Date: July 8, 2015
Primary Examiner: Prof. Dr.-Ing. Michael Beigl
Secondary Examiner: Prof. Dr. Moustafa Youssef

# Abstract of Dissertation

Human Activity Recognition (HAR) is essential for the interaction with Smart Environments. However, current sensors for capturing activities come with a number of drawbacks. In particular, sensors either require charging and attachment on a subject's body or an extended installation effort in order to cover the area of interest. In this dissertation the use of received signal strength (RSS) of IEEE 802.15.4 Smart Home radio technology was investigated as an alternative device-free sensing modality for HAR. Device-free operation means that no device has to be carried by the human for the recognition. Additionally, RSS is available in all consumer wireless technologies requiring no infrastructure investments.

Previously, device-free, radio-based research focused on localization or detection of subject presence. Recently, the investigations have been expanded to other topics including activity and gesture recognition, detection of breathing, among others.

This work excels over existing research as it facilitates RSS of low cost commodity transceivers for Activity Recognition. To demonstrate that RSS can be harnessed for this purpose we developed three theses which were validated in this dissertation:

1. RSS is influenced by human activities such that their recognition is feasible.
2. Parameters within and outside of the sensor system influence the quality of Activity Recognition.
3. It is possible to develop a practical online system for device-free, radio-based Activity Recognition.

The theses are tested and evaluated using empirical data stemming from activity experiments and simulation. Experiments were conducted in different IEEE 802.15.4 test beds. Feasibility was demonstrated using cross-validation with C4.5 decision trees, naive Bayes and k-Nearest Neighbour algorithms and comparing results to traditional accelerometer-based recognition. We found that the average accuracy of the device-free system was 86.8% which is comparable to the average accuracy of 85.3% for the accelerometer data. For the investigation of parameter influence we first derived the three aspects with corresponding parameters that determine how RSS is affected by a human activity. The experimental evaluation of five selected parameters using a Support Vector Machine classifier showed that their impact on performance decreases in the following order: size of environment (17%), radio transmission power (16%), frequency spacing (15%), frequency diversity (14%), length of a wireless link (11%). For a practical system for device-free HAR a number of challenges were identified. A holistic architecture comprised of four subsystems addressing these challenges was proposed and evaluated. These four subsystems are: 1) The Radio Sensor, which provides a highly synchronous sampling of the radio channel, 2) the LoS-Cross Inference System for the detection of subject presence, 3) the WiDisc Inference System for the differentiation of subjects, and 4) the RFHAR Inference System for Activity Recognition. Using a total of six measurement campaigns spanning a time frame of over one year it

was demonstrated that subject presence can be detected with 92% accuracy, even if a subject is not in motion; that three subjects can be discriminated with 67% accuracy when calibration data is derived from simulation; and that device-free Activity Recognition is possible with an average accuracy of 85% even 10 days after training.

Altogether, this dissertation comprises the following contributions:

1. Reference design of a device-free, 2.4GHz IEEE 802.15.4-based sensor system for Activity Recognition
2. Fundamental description of influences affecting Activity Recognition performance
3. Software design pattern for device-free, radio-based Inference Systems
4. Development and characterization of three specialized device-free, radio-based Inference Systems

# ZUSAMMENFASSUNG

Die Erkennung menschlicher Aktivitäten ist für die Interaktion mit intelligenten Umgebungen erforderlich. Die aktuell verfügbare Sensorik für die Erfassung von Aktivitäten weist eine Reihe von Nachteilen auf. So muss diese entweder regelmäßig aufgeladen und am Körper getragen werden oder erfordert einen hohen Installationsaufwand um die gesamte Umgebung abzudecken.

In dieser Arbeit wurde daher untersucht in wie weit die Signalstärke (RSS) von IEEE 802.15.4 Smart Home Funksendeempfängern als alternative gerätefreie Sensorik für die Erkennung von menschlichen Aktivitäten genutzt werden kann. Dabei bedeutet "gerätefrei", dass der Mensch für die Erkennung kein Gerät am Körper tragen muss. Da RSS einen Standardmesswert darstellt, der in Verbraucherhardware in der Regel vorhanden ist, benötigt dieser Ansatz keinen zusätzlichen Installationsaufwand, wenn bereits entsprechende Funkhardware vorhanden ist.

Seit 2007 beschäftigen sich verschiedene Forschergruppen mit der gerätefreien, funkbasierten Lokalisierung und der Präsenzerkennung von Personen. Erst vor Kurzem kamen neue Untersuchungen zur Aktivitäts- und Gestenerkennung oder auch zur Abschätzung der Atemfrequenz auf Basis von Funksignalen dazu.

Diese Arbeit geht über die bestehende Forschung hinaus, indem sie die RSS von sehr günstigen, für den Massenmarkt produzierten Funksendeempfängern für die Aktivitätserkennung untersucht. Für diese Untersuchungen wurden die folgenden drei Thesen formuliert:

1. RSS wird durch menschliche Aktivitäten derart beeinflusst, dass die Identifikation der Aktivität möglich ist.
2. Parameter innerhalb und außerhalb des Sensorsystems beeinflussen die Genauigkeit der Aktivitätserkennung.
3. Es ist möglich ein praxistaugliches Onlinesystem für die gerätefreie, funkbasierte Aktivitätserkennung zu entwickeln.

Die Thesen wurden schrittweise mittels empirischer Daten aus Experimenten und Simulationen validiert. Die Experimente wurden dabei in verschiedenen IEEE 802.15.4 Testumgebungen durchgeführt.

Die Machbarkeit der Erkennung (These 1) wurde durch Kreuzvalidierung mittels dreier typischer Klassifikationsalgorithmen (C4.5, Naive Bayes und k-nächste Nachbarn) gezeigt. Um die Ergebnisse bewerten zu können, wurde außerdem ein 3D Beschleunigungssensor eingesetzt. Dieser repräsentiert eine klassische Sensorik zur Aktivitätserkennung. Für die Daten des Beschleunigungssensors betrug die durchschnittliche Genauigkeit 85.3%. Für die Funkdaten wurden vergleichbare 86.8% erreicht. Damit wurde gezeigt, dass die Erkennung von Aktivitäten aus Funksignaländerungen prinzipiell möglich ist.

Um These 2 zu prüfen wurden zunächst die drei verschiedenen Aspekte identifiziert und spezifiziert die beeinflussen, wie sich eine menschliche Aktivität auf die RSS auswirkt. Diese Aspekte sind: Die Person, die Umgebung und der Funksensor, d.h. der Funksendeempfängeraufbau der die RSS misst. Anschließend wurden

fünf Parameter dieser Aspekte ausgewählt um sie mit unabhängigen Trainings- und Testdatensätzen und einem Support Vector Machine Klassifikationsalgorithmus hinsichtlich ihres Einflusses auf die Erkennungsgenauigkeit zu untersuchen. Die Parameter konnten wie folgt nach ihrem Einfluss geordnet werden: Größe der Umgebung (17%), Sendeleistung (16%), Abstand zwischen den Frequenzen, auf denen die RSS gemessen wird (15%), Anzahl der Frequenzen, auf denen gemessen wird (14%) und Distanz zwischen Sender und Empfänger (11%).

These 3 erforderte die Entwicklung eines praxistauglichen Systems für die Erkennung. Eine Analyse der bisherigen Untersuchungen und Anforderungen zeigte, dass dafür drei unterschiedliche Fragestellungen innerhalb eines ganzheitlichen Erkennersystems gelöst werden müssen: Die Erkennung, ob eine Person vorhanden ist, wer diese Person ist und welche Aktivität diese Person gerade durchführt. Um dies zu ermöglichen, wurde eine Systemarchitektur aus vier Komponenten entworfen: Der Funksensor, der für die Messung der RSS zuständig ist, das LoS-Cross System, das die Anwesenheit einer Person erkennt, das WiDisc System, das verschiedene Personen unterscheiden kann, und das RFHAR System welches die Aktivität der Person feststellt. Die drei Erkennersysteme wurden in ein Inferenzsystemensemble integriert, so dass diese die Ergebnisse der Erkennung gegenseitig nutzen können bzw. eine Erkennung nur dann durchführen wenn es sinnvoll ist (z.B. Erkennung einer Aktivität, erst wenn Person anwesend ist). Die einzelnen Erkennersysteme wurden mit insgesamt sechs verschiedenen Messreihen unterschiedlicher Komplexität getestet. Für die Umsetzung der Erkennersysteme wurde ein allgemeines Software Entwurfsmuster abgeleitet und beschrieben. Das LoS-Cross System erreichte eine Genauigkeit von 92%, wobei es dabei nicht erforderlich ist, dass sich die Person bewegt. Das WiDisc System erreichte eine Genauigkeit von 67% für die Unterscheidung von drei verschiedenen Personen. Es zeichnet sich inbesondere dadurch aus, dass es die Kalibrationsdaten aus Simulationen zur Ausbreitung elektromagnetischer Wellen ableitet. Das RFHAR System erreichte eine Genauigkeit von 85% für die Unterscheidung von drei verschiedenen Aktivitäten, sogar als die Kalibration schon 10 Tage zurücklag.

Die Prüfung der genannten Thesen führte zur Erarbeitung der folgenden Beiträge:

1. Referenzdesign eines gerätefreien, 2.4GHz IEEE 802.15.4-basierten Sensorsystems für die Aktivitätserkennung
2. Grundsätzliche Beschreibung der Einflussfaktoren für die Aktivitätserkennung
3. Software Entwurfsmuster für gerätefreie, funkbasierte Inferenzsysteme
4. Entwicklung und Charakterisierung von drei unterschiedlichen, spezialisierten Inferenzsystemen

# Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| ADL | Activities of daily living |
| AM | Amplitude modulation |
| C4.5 | Decision tree algorithm |
| CCA | Clear channel assessment |
| COTS | Commercial off-the-shelf |
| CSI | Channel state information |
| CSMA/CA | Carrier sense multiple access/collision avoidance |
| DECT | Digital enhanced cordless telecommunications |
| EM | Electromagnetic |
| ESD | Extreme studentized deviation |
| FM | Frequency modulation |
| GTS | Global time slot |
| IEEE | Institute of Electrical and Electronics Engineers |
| k-NN | k-Nearest Neighbors algorithm |
| LoS | Line of sight |
| MAC | Media access control |
| OFDM | Orthogonal frequency-division multiplexing |
| PCA | Principal component analysis |
| PCB | Printed circuit board |
| PIR | Passive infrared |
| RF | Radio frequency |
| RSS | Received signal strength |
| RTI | Radio tomographic imaging |
| SDR | Software-defined radio |
| SVM | Support vector machine |
| TDMA | Time division multiple access |
| TPR | True positive rate |
| WiFi | Wireless fidelity |
| WLAN | Wireless local area network |
| WPAN | Wireless personal area network |

# 1 INTRODUCTION

## 1.1 MOTIVATION

Today we are surrounded by an ever growing number of smart interconnected things. Examples include mobile phones, laptops, TVs, entertainment systems, microwave ovens, thermostats, fridges, among many others. In fact, in the Smart Home of tomorrow there will hardly exist a device or installation (lights, power outlets, switches) without network connection. Primary communication technologies of today's Smart Home are wireless and include WiFi, Bluetooth and low power mesh networks [17].

One of the main challenges in home automation is the comfortable control of these devices [54]. Approaches currently employed typically require explicit user input with classic interfaces such as personal computers, smart phones or tablet PCs. In these cases a device or application must be selected from a hierarchy of menus before a command can be issued. With increasing number of devices this type of control becomes more and more cumbersome, reducing user comfort and acceptance.

The growth prediction for the Smart Home market [4] indicates that the relevance of this challenge will further increase. The rise is further driven by demographic change. While an increase of elderly vs active workers puts significant pressure on the economy [15], it also lowers the chance that caregivers or family members are available to look after them. Providing intelligent surroundings which monitor the user and derive actions without active involvement is one way to enable a self-sustained living and avoiding the explosion of costs [56]. This paradigm is known as implicit interaction and describes "an action that is performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input" [68].

Such an implicitly controlled home automation system may be realized using two components: A sensor system component, which perceives the surroundings and derives the current context [68] and an actuation component, which uses this context and some (possibly learned) policy to act. Here, we understand context as "a situation and the environment a device or user is in" [69]. An important example of context for a Smart Home is the human physical activity [54]. In this work we will refer to human activities as activities of daily living as defined by Katz et al. [44] in 1963 and the process of Activity Recognition as follows:

- **Activities of Daily Living (ADL):** Activities which people perform habitually and universally.
- **(Human) Activity Recognition (HAR):** The process in which ADL are identified based on sensor data.

In the past 20 years a large corpus of research has been dedicated to Activity Recognition. Research has been very active in the domain of Pervasive Computing from which the vision of implicit interaction originates. Typical systems for Activity Recognition leverage body-worn sensors or infrastructure-based sensors [16]. Examples of body-worn sensors include accelerometers and gyroscopes, e.g. embedded in Smart Phones. Infrastructure-based sensors include camera systems, passive infrared (PIR) motion detectors or ultrasound. Both types of sensors typically have two major disadvantages:

1. **Installation Effort:** In order to establish the sensor system the corresponding hardware must be installed. This can mean that sensors must be installed in the users' home or that the user must be instrumented.
2. **Comfort of Use:** The user is regularly reminded of the presence of the system (e.g. visibility of sensors). Further on, the system may require additional actions from the user such as activating, charging or attaching the device to the body.

Creating a sensor system without such deficits to improve Smart Home sensing presents the main motivation of this work. An attached secondary motivation is to increase the understanding of the possibilities and limitations of the employed sensing concept.

## 1.2 GOAL AND SCOPE

One way to create a sensor system without the described drawbacks is by reusing information from existing infrastructures. In this case the user does not have to change his behavior and the sensing infrastructure is invisible as it merges with the existing installation. Previous research has considered this approach for Human Activity Recognition. Examples include the use of the house water infrastructure [25] and the electrical power consumption [28]. But these still have issues. Firstly, the exisiting infrastructure must still be adapted. However, the effort is reduced compared to a new installation. Secondly, they only allow the event-based recognition of user activities such as "toiletting" or "cooking". That is, activities that do not affect the infrastructure can not be recognized. In contrast, we seek an existing infrastructure that may be suited for sensing but that does not require explicit interaction with the installation. One promising example is the wireless communication infrastructure. Radio waves are ubiquitous, not perceived by humans and many Smart Home appliances are outfitted with a radio interface. Thus, a high level of comfort can be provided because subjects do not have to carry

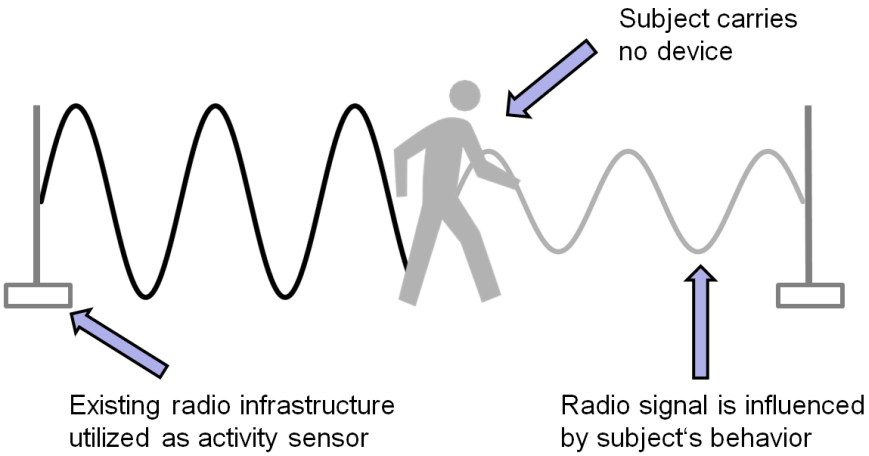



Figure 1.: The vision and concept of device-free, radio-based Activity Recognition

a device (**device-free**) and explicit infrastructure interactions are not required. As wireless devices further provide the means to measure signal quality, additional infrastructure adaptations can also be omitted.

**Therefore, the goal of this dissertation is to develop a novel sensor system that harnesses radio signal characteristics available in commerical off-the-shelf (COTS) radio hardware for device-free Activity Recognition.**

Figure 1 depicts the vision and the main aspects of this approach. The concept of the novel system stems from the observation that humans influence radio signals in a way that can be observed in COTS hardware [93]. As an illustrative example consider a kitchen AM radio: the quality and volume of the received station changes when a subject moves in the vicinity of the radio. **As an important limitation to this goal we define that we will only consider the impact of a single subject on the measurements over the course of this work.** The rationale is that increasing the number of subjects increases the complexity of the investigation. From a methodological point of view it is suggested to first analyse if the goal can be achieved using a simpler problem, before attempting to solve the more complex challenge.

## 1.3 THESES OF THIS DISSERTATION

In order to develop a sensor system for device-free, radio-based Activity Recognition using Smart Home wireless communication technologies we define three main theses which built on each other. They are verified in the described order over the course of this dissertation.

**Thesis 1** **Radio signals are influenced by human activities such that their recognition is feasible.**
This thesis focuses on the general feasibility of Activity Recognition using radio signal characteristics provided by Smart Home wireless communication technologies. It is verified in Chapter 3 by experimental evaluation, showing that human activities correlate with common radio signal characteristics.

**Thesis 2** **Parameters within and outside of the sensor system influence the quality of Activity Recognition.**
This thesis assumes that there are numerous environmental or system specific influences besides the actual activity which affect the recognition. In Chapter 4 we analytically decompose a typical measurement environment into general aspects. Each aspect has different parameters with varying influence on the recognition. A consecutive study on a parameter subset of two aspects provides experimental validation of the thesis.

**Thesis 3** **It is possible to develop a practical online system for device-free, radio-based Activity Recognition.**
This thesis assumes that a practical real-world recognition system can be constructed. Here, practical means that the information provided is useful and the system offers a certain robustness against environmental changes (e.g. environmental dispersion over time). In Chapter 5 we develop a suited holistic architecture which incorporates different subsystems of which each solves a specific challenge necessary for a practical recognition system. Each of these subsystems is designed, implemented and evaluated in Chapters 6–8 and Appendix A. The combination of the results is discussed in Chapter 9 thereby validating the thesis.

## 1.4 CONTRIBUTIONS

This dissertation has four main contributions. In a nutshell, contributions concern 1) the reference design for a practical device-free, radio-based HAR system; 2) the identification and description of parameters which affect system performance; 3) a software design pattern for device-free, radio-based context recognition systems (Inference Systems) and 4) the characterization of three such systems each focusing on a specific context. In the following these contributions and their potential for future research are briefly described.

**Contribution 1** **Reference design of a device-free, 2.4GHz IEEE 802.15.4-based sensor system for HAR**
The reference design has two main components: the Radio Sensor and the Inference System Ensemble. The Radio Sensor is a spatially distributed array of wireless transceivers with a custom firmware optimized for measuring human activities (40 Hz, 3 carrier frequencies). Measurements are consumed by the Inference System Ensemble conducting the HAR. The ensemble consists of three daisy chained Inference Systems: Presence Detection (LoS-Cross), Subject Discrimination (WiDisc) and Activity Recognition (RFHAR). The ensemble system output is a spatially constrained subject specific ADL.

This contribution provides future researchers with the understanding and tools to design, implement and build a practical Activity Recognition system

based on IEEE 802.15.4. It also provides the reasoning and description of the three required contexts for device-free, radio-based HAR.

**Contribution 2** **Fundamental description of influences affecting Activity Recognition performance**

This contribution describes the different aspects influencing recognition performance: subject, environment and Radio Sensor. Each of these aspects has a number of corresponding parameters such as the subject's location, the size of the environment or the transmission power, which are described. Five parameters are investigated in detail regarding their recognition influence using an experimental evaluation. The investigated parameters are ranked in respect to their influence.

This contribution supports future research as follows: Firstly, the parameter evaluation presents best practice guidelines for implementing device-free, radio-based HAR systems. Secondly, the decomposition and formalization of influences facilitates a structured methodology for new investigations. With limited effort it can also be adapted for investigating other contexts. Lastly, the parameter ranking highlights how the Radio Sensor may be further tuned for increased performance.

**Contribution 3** **Software design pattern for device-free, radio-based Inference Systems**

This contribution describes a general design pattern for device-free, radio-based Inference Systems. We find that such systems can be reduced to three modules: preprocessing, model generation and inference. The pattern describes the different possibilities of interconnecting these modules and their associated data. Each Inference System developed in this dissertation conforms to this pattern. Given a sample of recent systems from other researchers we further show that these can also be understood as implementations of this pattern.

Future researchers may find this contribution useful for the following purposes: to identify different implementation strategies for novel recognition systems; to analyze existing Inference Systems for alternative implementation approaches; and for evaluating existing implementations in respect to the employed design.

**Contribution 4** **Development and characterization of three specialized device-free, radio-based Inference Systems**

We show that to provide practical device-free, radio-based HAR three challenges must be addressed: Presence Detection, Subject Discrimination and Activity Recognition. For each of these challenges we developed a specific Inference System. Each system has a number of parameters which we analyze in extensive evaluations. These encompass different preprocessing configurations (e.g. filters, aggregations), feature configurations (feature types, feature window/shift), calibration methods (3D modelling and simulation,

groundtruth annotation) and classification algorithms (machine learning algorithms, novel recognition scheme), among others.

This contribution fosters future research by providing novel algorithms and methods for device-free, radio-based HAR which may also be applied to other contexts/sensors and by highlighting research opportunities through the characterisation and discussion of the presented systems.

## 1.5 STRUCTURE OF THIS DISSERTATION

Figure 2 shows the structure of the work highlighting the chapters which confirm the theses. **Chapter 1 and 2** motivate this work, describe background details and report on the related work and the difference to this dissertation. **Chapter 3 to Chapter 9** present the core of the thesis. In **Chapter 3** we investigate the feasibility of radio-based Activity Recognition using experimental studies. This chapter proves the validity of Thesis 1. In **Chapter 4** the effects of parameters of the envisioned sensor system are quantified. For this purpose we first identify typical parameters and select the most potential ones for further experimental evaluation. We show that indeed they have significant influence on recognition performance (up to 17%) confirming Thesis 2.

Using the insights from Chapter 3 and Chapter 4, we derive a complex architecture for a practical device-free Activity Recognition system in **Chapter 5**. As detailed in this chapter, implementing a robust and useful device-free, radio-based HAR system requires first to detect a subject, then discern its identity and finally recognize her activities. Hence, together with the overall discussion in Chapter 9 it creates the frame for the validation of Thesis 3. In this frame the systems addressing the mentioned challenges (Chapters 6-8) and the Radio Sensor implementation (Appendix A) are linked together into a holistic system. **Chapter 6** describes LoS-Cross a system which uses a special transceiver topology to detect subject presence even if the subject is not moving. In **Chapter 7** we develop WiDisc. WiDisc uses 3D modelling and simulation instead of cumbersome manual fingerprinting to determine the class of a subject. In **Chapter 8** we present RFHAR, which is the actual Activity Recognition system. RFHAR is optimized and tested on a large number of data sets. The system is also evaluated with respect to training data age and amount. It is shown that recognition accuracy is only slightly affected when minimizing the training data. In order to conduct real world experiments a radio sensor system was constructed. Throughout this work we refer to this system as Radio Sensor. The implementation of the Radio Sensor using COTS IEEE 802.15.4 transceiver modules is described in **Appendix A**. Finally, in **Chapter 9** the results of the presented inference systems are discussed in the context of the proposed complex architecture. The outcome of the discussion in combination with the actual evaluation results presents the proof for Thesis 3. The dissertation concludes with **Chapter 10**.

Contribution 1 is provided through the results of **Chapters 4–9** and **Appendix A**. Contribution 2 stems from the analysis and evaluation of parameters in **Chapter 4**.

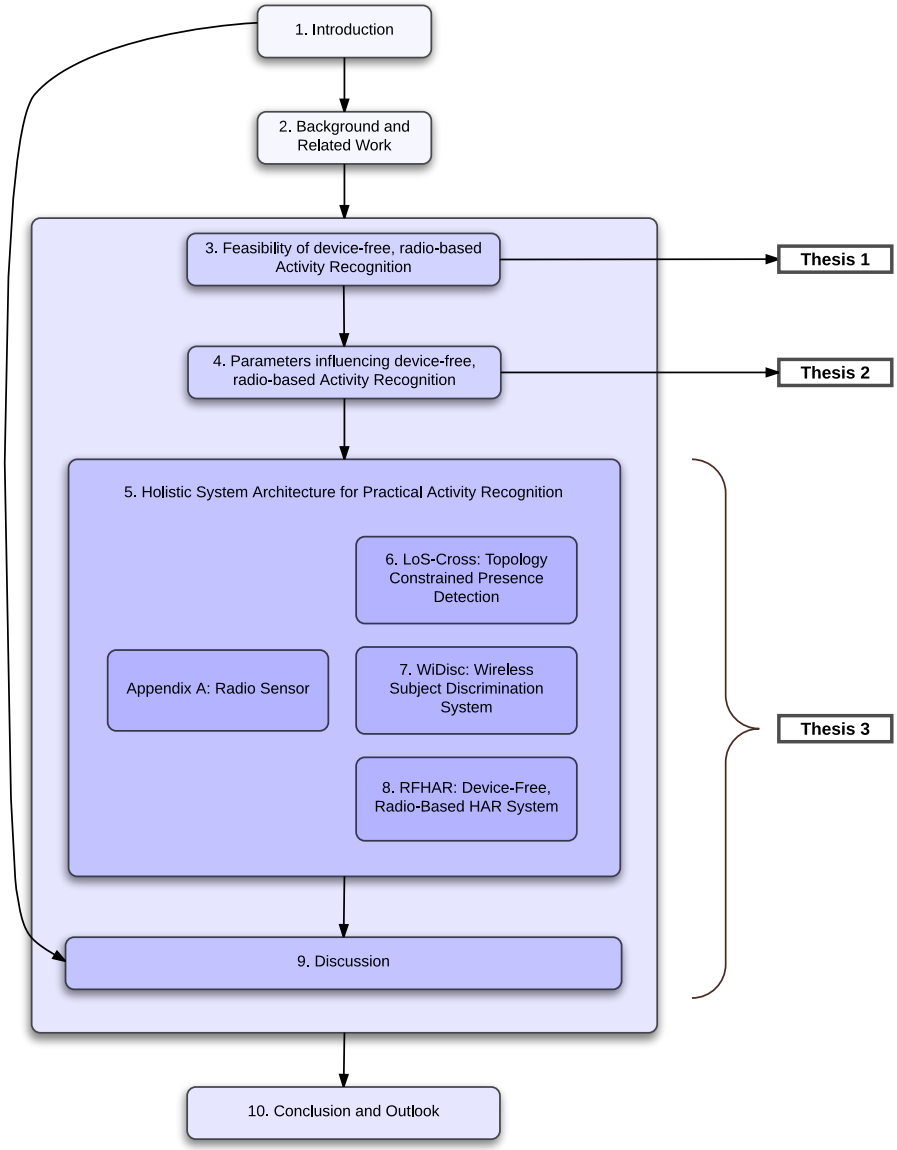



Figure 2.: Structure of the dissertation

Contribution 3 is provided by the design consideration in chapter **Chapter 5** and the discussion in **Chapter 9**. Finally, Contribution 4 is provided by the description and evaluation of the Inference Systems in **Chapters 6–8**.

# 2 Background and Related Work

This chapter provides a brief overview over the background and related work of this dissertation. In addition, most chapters come with a related work section of their own which focuses on the specific challenge. The chapter closes with a summary highlighting the gap in current research with respect to this work.

## 2.1 Background

### 2.1.1 *Electromagnetic Propagation*

Parts of this section are based on our paper [75]. Radio waves are electromagnetic waves with a frequency ranging from 3 kHz to 300 GHz. The propagation of electromagnetic waves originating from an ideal omnidirectional antenna in free space is described by the free space path loss equation (FSPL) [63]:

$$\mathrm{FSPL}[\mathrm{dB}] = 10 \log_{10} \left( \frac{4\pi}{c} \times d \times f \right) \tag{1}$$

Here, $c$ is the speed of light, $d$ is the distance from the transmitter and $f$ is the carrier frequency. The received energy falls in proportion to the square of the distance from the transmitter and the signal's frequency (aka inverse square law).

Furthermore, electromagnetic waves are susceptible to additional effects when propagating. The character of these effects is mainly a function of frequency, transmission media and objects encountered during propagation. Such effects include reflection (wave partially bounces off an object), refraction (change of direction when passing from one medium to another), absorption (dissipation of energy on interaction with an object), diffraction (wave is bend around an obstacle), scattering (wave bounces off in multiple directions) and polarization (change in orientation of wave oscillations upon interaction) [63].

Another aspect of radio wave propagation is multipath propagation. Typically, the transmitter's antenna will emit radio waves in all (omnidirectional antenna) or in specific directions (directional antenna). Also some of the above enlisted effects occur simultaneously. For instance, a radio wave may propagate through an object but part of it will be reflected on its surface and some of its energy is

absorbed by the object. After this interaction there are (at least) two radio signals propagating in different directions but originating from the same source. Radio signals originating from the same source, which reach the receiver by two or more paths, are called multipath signals or components. The power of the received signal is the sum of these destructive or constructive multipath components. More formally the received power may be expressed as:

$$P_{RX}[W] = \left| \sum_{i=0}^{N-1} a_i exp(j\theta_i) \right|^2 \tag{2}$$

Here $P_{RX}$ is the instantaneous power in the receiver. $N$ is the number of multipath components or rays. $a_i$ is the real amplitude of the $i$th multipath component and $\theta_i$ is the phase of the $i$th multipath component. In COTS hardware, there is usually a single parameter available which measures the received signal's power: the received signal strength (RSS) which is typically a scalar value. Based on Eq. 2 RSS can be understood as the result of a complex function of the above described effects over the course of the signals through space until entering the antenna of the receiver.

### 2.1.2 *Metrics for evaluating Inference Algorithms*

Throughout the thesis we will evaluate a number of algorithms and algorithm configurations which determine a specific contextual information from input data. In order to benchmark their performance a metric is needed. The most common metric is the accuracy.

Equation 3 shows the computation of the accuracy for multi class problems [85] when training a classifier with a training set $t$ and testing with a test set $s$. Therein $C$ refers to the considered classes, $i$ is the current class for which the tp, tn, fp and fn are the number of true positives, true negatives, false positives and false negatives, respectively.

$$\text{Accuracy}(t,s) = \frac{\sum_i^C \frac{\text{tp}_i + \text{tn}_i}{\text{tp}_i + \text{fn}_i + \text{fp}_i + \text{tn}_i}}{C} \tag{3}$$

Accuracy may be flawed when the number of samples per class in the data set is not identical, i.e. if the classes are not balanced. Other metrics such as F-Measure provide an alternative in such cases. However, even with F-Measure the effect of an overrepresented class in the training data set can not be evaluated easily. For this reason, classes in all data sets employed in this work were balanced prior to evaluation. For some investigations we use confusion matrices [86] to analyse the result. By the means of the true positive rate (TPR) they provide an indication how well a specific class is identified by the particular algorithm.

The true positive rate (TPR), also referred to as sensitivity, describes the relation of correctly classified true positives (tps) to the total number of positives (ps) for this class $i$.

$$\mathrm{TPR}_i = \mathrm{sensitivity}_i = \frac{\mathrm{tp}_i}{\mathrm{p}_i} \tag{4}$$

## 2.2 RELATED WORK

### 2.2.1 *Traditional Activity Recognition Research*

The recognition of human activities is tied to the fundamental vision of Pervasive Computing. If actions of a subject can be recognized and predicted, the door is opened to a variety of applications improving the quality of life. Analysis of physical activities originates from medical or physiological research in which activity execution was evaluated to judge the performance or health status of individuals, e.g. [13]. On the other hand, Activity Recognition was studied in computer vision since the late 90s (cf. [96]). But the basis for the current state of Activity Recognition in the field of Pervasive Computing was laid in 2004 by Bao and Intille [8]. They used five 2D accelerometers on each of 20 subjects of different age groups and achieved up to 84.29% recognition accuracy for a large number of activities of daily living. This was made possible by what is known as the tool chain of Activity Recognition today: gather sensor data, annotate data, split data, compute features, train and evaluate. Since 2004, the Pervasive Computing community has conducted a multitude of investigations utilizing the original tool chain. Nowadays, the recognition of activities of daily living using accelerometers is considered well investigated in the research community and only marginal advances are made. For instance, He et al. [30] presented a novel algorithm in 2012 increasing the accuracy for an activity dataset similar to the one of Bao et al. to 99%. Open challenges include enabling Activity Recognition without elaborate training [9], novel application cases [70], using new sensors (this work) or combining existing sensors for better recognition or even deriving the activity of a whole group [26]. For further information on the field, the reader is referred to [8] and [16].

### 2.2.2 *Device-Free Radio-based Context Recognition*

Device-free, radio-based context recognition is a highly active field of research continuously expanding in new directions. The following subsections provide an overview over the current research limited to the most prominent publications. At the end of the section we further provide an overview of the most active research groups in this domain.

*Scope, Devices and Signal Information*

Based on our vision, we consider research related to this work only if the employed technology is originally intended for narrowband communication. Thus, we do not consider radar-based systems or ultra-wideband systems.

Today, related research is typically conducted on either one of three different hardware devices which provide different types of signal information:

- Software Defined Radios (SDR): Provide the highest flexibility, bandwidth, sampling frequency and resolution. Measured characteristics are amplitude and phase of a radio channel.

- WiFi or other COTS transceivers: Typically measure received signal strength (RSS) per packet. RSS is a low resolution value represented in a single byte.

- Modified 5300 Intel Network Interface Controller (NIC): Measures Channel state information (CSI) per packet. A CSI measurement contains the amplitude and phase of the 30 IEEE 802.11n orthogonal frequency-division multiplexed (OFDM) subcarriers.

The possibility of measuring CSI became available only recently. It can be measured if a specific firmware is uploaded to the NIC [29]. Further on, the measurements are only carried out for a special type of packet and in an unencrypted WiFi network. For these reasons, we do not consider it common Smart Home technology. However, CSI provides much more information (compared to RSS) at relatively low cost and a manageable amount of data (compared to SDR). For this reason, researchers have been able to achieve impressive results and the interest in CSI is further growing.

*Initial Experiments*

In 2006, Woyach et al. [93] conducted RSS-based experiments and found that the localization of objects in-between nodes without wireless connectivity could be feasible. They observed a difference in RSS changes if an object moved between transmitter and receiver vs when it moved only in the vicinity. They identified the RSS variance as a feature to extract further insights into the type of movement. Further investigations showed that variance differed depending on the trajectory of the object, the wireless network topology and the environment. They also showed that the change in position of transceivers has a much stronger impact on RSS than the movement of an external object. The same experiments were also used to show a radio wave property to which Woyach et al. refer as spatial memory. Therein they showed that after any kind of temporary change in the environment the RSS returned to the initial values. These first experiments provided the ground for a large corpus of device-free, radio-based research described in the following sections.

*Presence Detection and Localization*

In 2007, researchers [99, 101] demonstrated the feasibility of using RSS to derive the presence and the location of a subject not carrying a wireless device. Since then especially the development of RSS-based radio tomographic imaging (RTI) [92] has influenced device-free localization research strongly. Until today algorithms related to this method have shown localization errors as low as 10 cm [41] for classic indoor and 30 cm [42] for through-wall scenarios. However, RTI usually uses a large number of low power transceivers. Thus, other researchers have considered robust presence detection [45] and localization [66] using WiFi, achieving accuracies as low as 1.7 m and 2.2 m for single and multiple subjects, respectively. CSI-based localization has also been investigated [1] showing superior performance using fewer links. When using few nodes subject localization is typically based on fingerprinting which requires costly calibration measurements. An alternative has been presented recently by Heba and Youssef [6] who attempted localization using simulation generated RSS fingerprints.

*Activity Recognition*

In 2011, our group [74] for the first time presented an online system employing two SDR which could identify two different kinds of activities: walking and talking on the phone. The system further recognized the state of the room's door. It achieved an overall accuracy of 88%. The recognition of additional ADL was further demonstrated together with Sigg et al. [82] using cross validation on SDR. The achieved classification performance was 82% for four activities conducted at 11 locations. Shi et al. [78] showed the feasibility of ADL recognition using FM radio signals analyzed using SDR. In 2013, our group [73] showed the feasibility of ADL recognition using COTS radio modules (parts of this study are presented in the next chapter). Recently, Wang et al. [90] demonstrated the recognition of ADL using CSI achieving accuracies of 90% and 97% in two test beds with independent training and test sets for 7 activities.

*Gesture Recognition*

In Ref. [61] Pu et al. reported an accuracy of 97% for the recognition of eight gestures using SDR. Sigg et al [81] showed, albeit with reduced recognition accuracy, that this is also possible using WiFi RSS measured using standard mobile phones. Recently, Abdelnasser et al. presented a similar RSS-based gesture recognition system which achieved an accuracy of up to 96% without calibration [2].

*Further Contexts*

In Ref. [5] Al-Husseiny et al. presented techniques to distinguish cars from humans using WiFi RSS measurements. In Ref. [58] CSI was utilized to localize a subject based on her breathing motion, while at the same time determining the breathing rate. Earlier, the group also demonstrated breathing rate estimation using RSS

measured in many transceivers [60]. In 2015, Abdelnasser et al. showed that breathing estimation is also possible using WiFi RSS [31]. In Ref. [95] authors demonstrated counting of up to four subjects with an average accuracy of 86% within large indoor areas (150 $m^2$ and 400 $m^2$). In Ref. [51] researchers showed that they can detect subject falls using a 2-level transceiver topology with 100% accuracy. In Ref. [49] authors used channel state information to determine sleep postures and respiration status. In Ref. [89], the authors used software defined radios to classify different words based on signal distortions due to mouth motion. In Ref. [34] Huang et al. explored the possibility to create 2D images by moving a SDR in front of an object and extracting reflected WiFi signals from the radio channel.

*Most active Research Labs*

At the time of writing there are a number of research labs and individual researchers who have continuously worked on device-free, radio-based recognition over the past years and whose contributions make up most of the related work.

- Neal Patwari's group at the SPAN Lab at Utah University: Presence detection, localization and tracking using a large number of COTS transceivers and detection of breathing rate using RSS [60] and CSI [58].
- Moustafa Youssef's group at the Wireless Research Center at the Egypt-Japan University Alexandria: Presence Detection [45], localization and tracking using WiFi with and without channel state information [66, 1]. Further research involves object differentiation [5], gesture recognition [2] and breathing estimation [31].
- Stephan Sigg at the University of Göttingen: Activity recognition with focus on SDR [82] and gesture recognition using mobile phone WiFi RSS [81].
- Lionel M. Ni's group at the University of Science and Technology Hong Kong: Motion detection [94], localization and tracking using many COTS transceivers [100]. Speech recognition using channel state information [89].

However, as attention for device-free context recognition is constantly increasing this list is a snapshot of the active research community at the time of writing (May 2015).

## 2.3 SUMMARY

In this chapter we presented a cross section of related work and basic background material. The previewed research does not proof any of the defined theses as it uses either different hardware and measurement quantities or is concerned with the recognition of different contexts.

However, it still offers important insights regarding methods and approaches: Generally, the positive results of SDR and CSI investigations as well as work on

Presence Detection using RSS indicate that Activity Recognition using RSS can be possible (Thesis 1). The related work also provides insights for the planning of experimental investigations: as RSS was shown to be strongly influenced by location this should be considered in activity experiments. Further on, early research by Woyach et al. and Youssef et al. showed that signal dispersion is a stable and insightful feature for motion characterisation. Localization research was further concerned with the creation of parameterizable models. Parameters of these models could be potential candidates to investigate for Thesis 2. Finally, the traditional Activity Recognition tool chain provides a well tested basis for the implementation of a practical online system (Thesis 3).

# 3 Feasibility of device-free, radio-based Activity Recognition

In this chapter we show the general feasibility of Activity Recognition from received signal strength measurements (Thesis 1).

The chapter starts by pointing out why Thesis 1 is not yet validated albeit SDR-based Activity Recognition studies have been previously conducted. To provide this validation we conduct an experimental study which is described in the next section. Thereafter the evaluation of the data set is presented. The chapter is closed with a discussion of the findings and a conclusion.

## 3.1 Introduction

In this chapter we investigate if human activities are reflected in the received signal strength (RSS) of commercial of the shelf (COTS) wireless transceivers. Prior to the publication of our paper [73] in 2013, on which this chapter is based, device-free Activity Recognition was only investigated using SDR (cf. Chapter 2). These devices differ significantly from consumer COTS radio hardware with respect to the following properties: channel sampling rate (SDR: 256 kHz vs COTS: 40 Hz), channel bandwidth (flexible (8 MHz-256 KHz) vs fixed (2 MHz)), available radio signal information (Phase, Amplitude, Frequency (I/Q values) vs RSS), signal resolution (complex float vs single byte RSS), receiver sensitivity, among others. For these reasons findings based on SDR hardware are not sufficient to support Thesis 1 of this dissertation. Hence, we conducted a feasibility study employing exclusively COTS transceivers. To compare the RSS-based measurements to a classic Activity Recognition sensor, the subject also carried an accelerometer during the experiments. For analysis we utilized 10-fold cross-validation with three machine learning algorithms commonly used in Activity Recognition. This approach is based on the assumption that if there is a correlation between human physical activities and RSS measurements, activities will incur characteristic patterns in the data which can be discovered by a suited pattern recognition algorithm. The results indicate that Activity Recognition is feasible with COTS transceivers and that the achieved performance is comparable to the classical approach (ca. 90% accuracy).

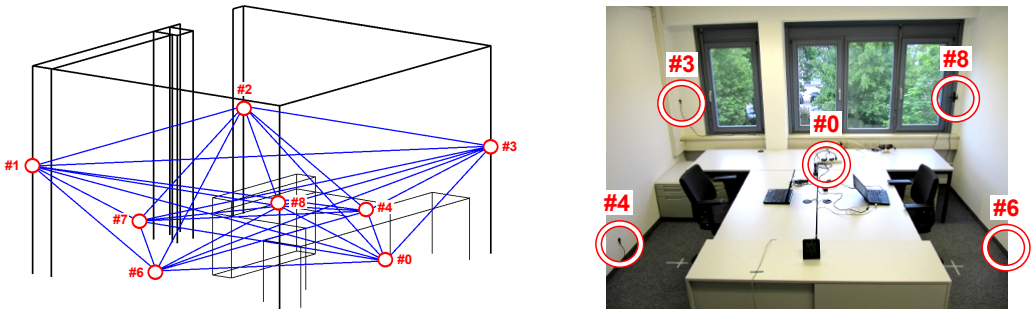



Figure 3.: Radio Sensor Installation

## 3.2 DATA COLLECTION

In this section we describe the two primary aspects concerning data collection: the test bed installation and the conducted activity experiments.

### 3.2.1 *Test bed*

The test bed consisted of a number of COTS radio transceivers installed in an office room and the accelerometer attached to the subject. As described in detail in Chapter 4, we refer to a configuration of transceivers as Radio Sensor. The sensor hardware and measurement algorithm is described in Appendix A, it allows to take RSS measurements at 40 Hz. The Radio Sensor implementation in this test bed used a single carrier frequency ($f_c = 2475\,\text{MHz}$) and 8 transceivers outfitted with omnidirectional ceramic antennas. In addition to the Radio Sensor installation, the subject carried a mobile sensor node with an ADXL335 3D accelerometer which was also sampled at 40 Hz. The room was 4.04x5.33 m, with gypsum walls on three sides and a window front. It had a bare concrete ceiling at the height of 2.70 m. The room featured typical pressboard furniture, two office chairs and two laptop computers. We deployed 8 nodes in the experiment room in different heights to consider upper and lower body halves separately and thereby increase accuracy [8]. The location of the nodes and the wireless links are shown in Figure 3. We chose 1.40 m, approx. the height of the human torso and impacted by arm movement, for the nodes #1, #2, #3, #8 which are deployed in the corners. We chose 0.30m, approx. the height of the middle of an adult's shin, for the nodes #0, #4, #6, #7 located in the middle of the walls. Thereby node distance was between 2m to 6m, which has been shown to work for device-free RSS-based human motion detection [102]. To compare the performance of the device-free recognition to the typical device-bound approach the test bed also included a mobile node (#5) with an accelerometer. Bao et al. [8] showed that an ideal location for recognizing a multitude of activities using a single sensor is the user's hip. Therefore, node #5 was attached to the hip of the subject.

### 3.2.2 *Experiment*

During the experiment a subject performed a range of activities which were recorded using the test bed installation. We selected activities of daily living, as well as activities that have been frequently investigated using accelerometers and such that seem hard to discriminate using accelerations. The selected activities were "Walking", "Standing", "Sitting", "Sitting and Typing", "Lying", "Lying and Waving" and "Outside (the room)". Existing research showed that localization using RSS is feasible [75]. Hence, we needed to ensure that indeed the activity and not the location of the person is recognized. Therefore the in-place activities "Standing", "Lying" and "Lying and Waving" were performed at five different locations (A-E) inside the room (capital letters in Figure 4). Activities were conducted for a total of 40 s each as follows:

1. Outside room (20 s)
2. Walking randomly across all locations (total time: 40 s)
3. Standing at each location for 8 s
4. Sitting at location D and E for 20 s each
5. Sitting and Typing at location D and E for 20 s each
6. Lying at each location for 8 s
7. Lying and Waving at each location for 8 s
8. Outside room (20 s)

Activities were conducted by each of the two subjects while the room door was either open, half open or closed. Thus, each sequence was repeated 6 times. Recordings were conducted in three sessions on three consecutive working days. Sessions lasted 1-2 hours and took place at 10pm, 3pm and 5pm. Subjects conducting the activities were male, 1.75m, 85kg and male, 1.76m, 72kg. The total recording time for each activity was $6 \times 40\text{s} = 4\text{min}$. Since we considered 7 different contexts, the total length of the data set was 28min.

## 3.3 EVALUATION

In this section we describe the conducted preprocessing, the evaluation methodology and the results of the evaluation after applying typical machine learning algorithms.

### 3.3.1 *Preprocessing*

The raw data of the experiments consisted of approx. 70,000 samples. Each sample had 59 attributes: the RSS for all wireless links in both directions (56 single way links for 8 transceivers) and the 3-axis acceleration information occurring at node #5 at the subject's waist. Recent studies showed that asymmetry between two way link measurements is primarily an effect of transceiver power and receiver sensitivity [52]. Hence, we averaged the RSS of the two way links, leaving 28

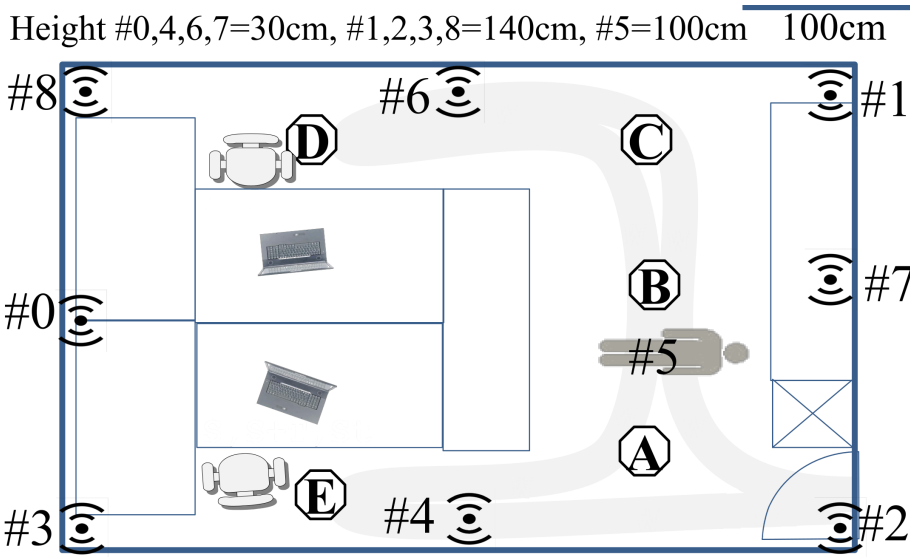



Figure 4.: Top down view of the experiment room with transceiver positions (#), activity locations (◯) and a typical walking path (gray).

two-way link attributes. Zero RSS values which indicated lost packets were not removed from the data. Based on literature in classic Activity Recognition [8] and device-free RSS-based localization [102] mean and variance over non-overlapping windows of 40 data instances (=1s) were selected as features. We chose non-overlapping windows to avoid including training information in test data when performing 10-fold cross validation. For evaluation, the data was split in two data sets based on the used sensor:

- Device-bound accelerometer data with 6 features: mean and variance for each axis of the accelerometer from the mobile node #5.
- Device-free Radio Sensor measurement data with 56 features: mean and variance for all 28 links of the Radio Sensor.

Finally, the "Outside" (the room) class was excluded from the accelerometer data.

### 3.3.2 *Examination of Raw Data*

Based on the raw data (Figure 5) most of the activities look significantly different in both sensors. For the accelerometer our expectation was that activities with similar poses would show similar signal patterns. But for the activities "Sitting and Typing" vs "Sitting" and "Lying and Waving" vs "Lying" the observed measurements did not support this assumption. Examining the video capture of the experiments showed that in both cases a body posture change takes place which changes the acceleration measurements strongly: "Sitting" is performed leaned back while "Sitting and Typing" is performed bowed forward to reach the keyboard. Likewise

"Waving and Lying" requires a more upright body posture in order to move one the arms compared to "Lying".

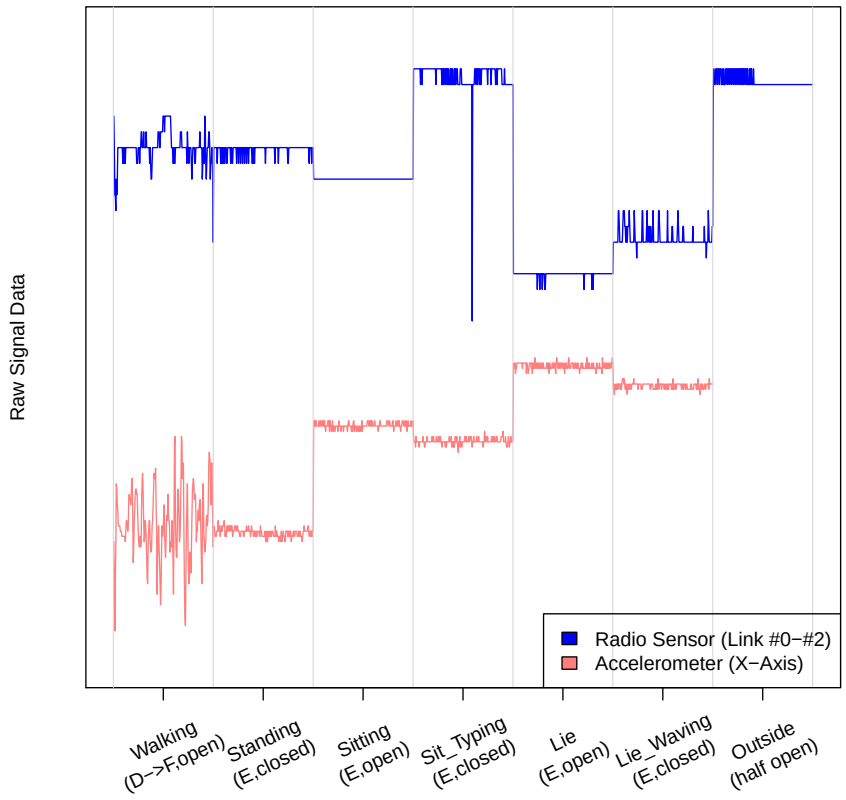



Figure 5.: 5s raw data for each investigated context from both sensors

In the accelerometer data the walking activity showed a very high variance. In contrast, all other activities cause similar fluctuations and could only be distinguished by their absolute values. Similarly, the Radio Sensor showed the strongest RSS fluctuations for the walking activity. In addition, some of the other activities did not only differ in their absolute signal levels but also in their variances (note esp. the difference between "Waving" and "Sitting and Typing").

We speculate that the magnitude of RSS fluctuations correlates with the size of the object being moved in the link. However, it is also likely a function of the subject's location and her volume. E.g. "Standing" and "Typing" produce similar signal changes. Another reason for peak signal fluctuations are packet losses, which are not handled in this early Radio Sensor implementation (e.g. down peak in the figure for "Sitting and Typing"). Nevertheless, this brief examination suggested that Activity Recognition using RSS COTS transceiver information could be feasible as all activities have a specific signature in the measurements.

### 3.3.3 *Distinguishing Activities using Pattern Recognition*

In this section we present the evaluation of the collected data set using typical machine learning algorithms with 10-fold cross-validation. In order to avoid classifier dependent results [27] three common Activity Recognition classifiers were selected for evaluation: k-Nearest Neighbours (k-NN) with k=10, naive Bayes and C4.5 decision tree. We used the implementation of these classifiers in the Orange data mining framework [21]. The classifiers were trained on 9/10th of the data and tested against the remaining 1/10 until all deciles of the set were tested (10-fold cross-validation). Figure 6 shows the average over these 10 iterations for both sensors. The multi class accuracy was used as metric (cf. Equation 3).

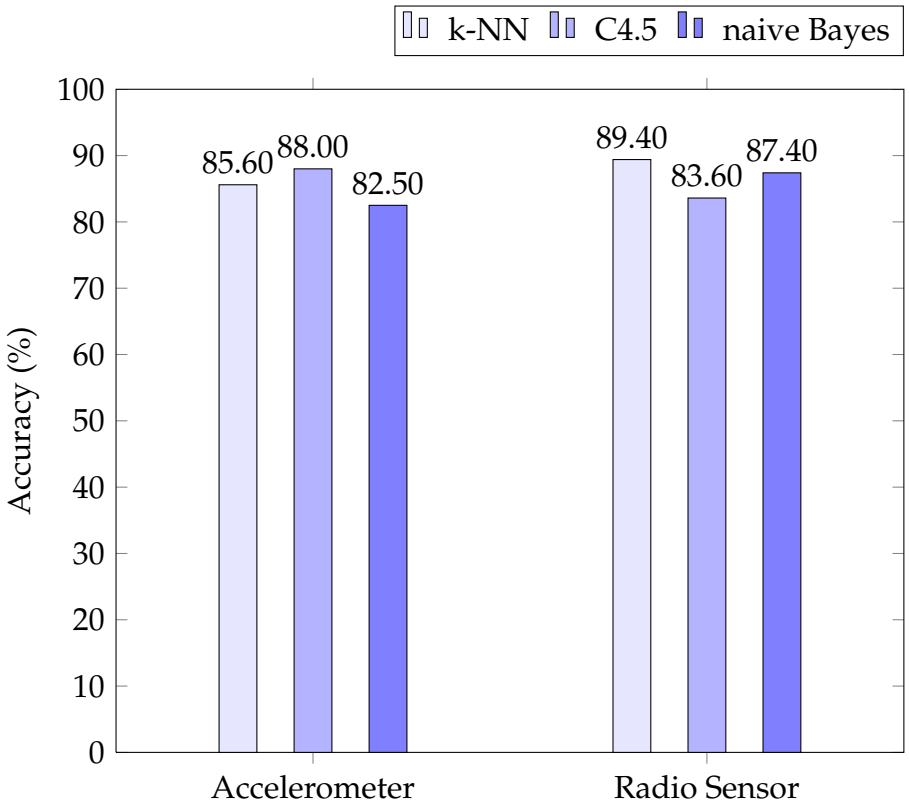



Figure 6.: Cross-validation accuracy across classifiers for both sensors

The best result for the accelerometer data was 88.0% using the C4.5 algorithm (Fig. 6). The best result for the Radio Sensor was 89.4% using the k-NN algorithm. A reason for the superior performance of k-NN in the Radio Sensor may lie in its distance based classification. As the distance computation gives equal importance to all features both in training and testing, a feature which looked promising in the training set is not favored when testing, perhaps preventing overfitting.

In contrast, C4.5 might use a feature providing good discrimination in the training set as top level decision feature failing on the test data. Average accuracy for accelerometer data was 85.3%, while average accuracy for the Radio Sensor was 86.8%. Also note that the performance of the Radio Sensor classification may

be considered superior as it has an additional class (Outside) which was excluded from the accelerometer data.

**The results support Thesis 1 in two regards:**

1. **There is a correlation between RSS and the conducted activities which allows to differentiate between them.**

2. **The achieved performance is comparable (or superior if considering the additional class) to classic Activity Recognition sensors.**

To gain insights in classification differences between sensors we analysed the results in more detail.

| Truth \ Predicted | Walking | Standing | Sitting | Sitting+Typing | Lying | Lying+Waving |
|---|---|---|---|---|---|---|
| Walking | 0.92 | 0.07 | 0 | 0 | 0 | 0.01 |
| Standing | 0.05 | 0.95 | 0 | 0 | 0 | 0 |
| Sitting | 0.01 | 0 | 0.96 | 0.03 | 0 | 0 |
| Sitting+Typing | 0 | 0 | 0.05 | 0.95 | 0 | 0 |
| Lying | 0 | 0 | 0 | 0 | 0.72 | 0.28 |
| Lying+Waving | 0 | 0 | 0 | 0 | 0.39 | 0.61 |

Table 1.: Confusion Matrix for the C4.5 classifier on the accelerometer data

| Truth \ Predicted | Walking | Standing | Sitting | Sitting+Typing | Lying | Lying+Waving | Outside |
|---|---|---|---|---|---|---|---|
| Walking | 0.71 | 0.08 | 0.03 | 0 | 0 | 0.04 | 0.14 |
| Standing | 0.05 | 0.95 | 0 | 0 | 0 | 0 | 0 |
| Sitting | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Sitting+Typing | 0 | 0 | 0.04 | 0.96 | 0 | 0 | 0 |
| Lying | 0 | 0 | 0 | 0 | 0.92 | 0.07 | 0.01 |
| Lying+Waving | 0 | 0 | 0 | 0 | 0.16 | 0.83 | 0.01 |
| Outside | 0.01 | 0.02 | 0.02 | 0 | 0.05 | 0.02 | 0.88 |

Table 2.: Confusion Matrix for the k-NN on the Radio Sensor data

The confusion matrices for the best classifiers for the accelerometer (Tab. 1) and the Radio Sensor (Tab. 2) indicate the following:

- "Walking" was best discriminated using the accelerometer data (92% true positive rate vs 71%). It was the activity with the smallest true positive rate in the Radio Sensor.

- "Standing" and "Sitting" were both discriminated with comparable performance (>95%) in both data sets. However, the Radio Sensor achieved slightly better results as it has less confusions with "Sitting and Typing". In the accelerometer data confusions of "Sitting and Typing" vs "Sitting" were expected as the activities are very similar.

- "Lie" and "Lying and Waving" are confused much more often using the accelerometer data than the Radio Sensor (up to 20% difference). This was expected, because the raw data patterns from the accelerometer measurements were very similar (Fig. 5).

- Most misclassifications in the Radio Sensor data set originate from confusions of the "Outside" (the room) activity with any of the other activities

## 3.4 DISCUSSION

### 3.4.1 *Data Collection Setup*

The achieved results indicate the feasibility of Activity Recognition using RSS measurements. However, the study was conducted using a single specific experiment setup involving a large number of transceivers, specific transceiver configurations, topology and room geometry. While a number of parameters varied in the experiments (door state, subjects, time of day) it is not clear at this point how other parameters influence recognition performance. In order to develop online recognition systems, parameter influences must be investigated. For this reason, we conducted a number of parameter investigations described in Chapter 4. In the long-term, understanding parameter influences may allow the creation of robust recognition models which can be parameterized.

### 3.4.2 *Evaluation Methodology and Results*

Evaluation was conducted using cross-validation using a number of typical classification algorithms. Cross-validation is prone to overfitting and typically leads to overconfident results. Nevertheless, cross-validation can be employed to fathom the upper bound of the possible recognition performance [10]. Therefore, we concluded that indeed patterns were present in the data which are unique for the conducted activities. How well these patterns can be discovered in an online trial, should be investigated as next step utilizing independent test and training sets. Further on, the experiments showed that the Radio Sensor performance was reduced by confusions with the outside room context. We address both of these challenges starting from Chapter 5 in which we derive the architecture of a practical, online device-free Activity Recognition system.

## 3.5 CONCLUSION

The described study shows the validity of the first thesis: device-free Activity Recognition using simple, low-cost commercial off the shelf transceivers is feasible. Indeed, with the conducted study we showed that achievable cross-validation performance is comparable to typical Activity Recognition sensors (accelerometer). However, feasibility is only the first step towards an actual practical recognition system. Hence, in the following chapters we investigate parameters that influence the recognition performance and iteratively develop a practical device-free online system for Activity Recognition.

# 4 Parameters Influencing Device-Free, Radio-based Activity Recognition

This chapter describes the validation of Thesis 2, showing that parameters within and outside of the sensor system influence Activity Recognition quality. The chapter is structured as follows: In the introduction we identify three different aspects affecting RSS changes in the presence of an ADL. Next, we review related research for similar investigations. Thereafter, the analysis section defines parameters associated with the mentioned aspects in more detail. Using this definition an evaluation approach and guidelines for estimating the impact of parameters is developed. In the experiment section we report from a total of five parameter investigations and rank the parameters based on their impact. The chapter is closed with a discussion and conclusion.

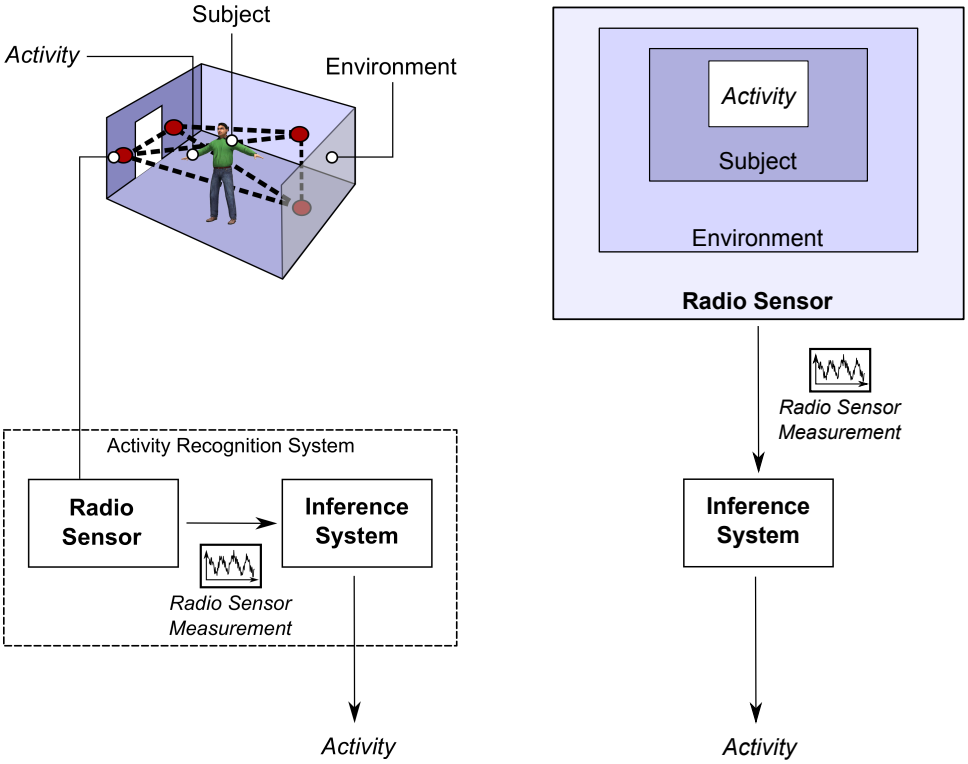
## 4.1 Introduction

A device-free, radio-based sensor system is typically comprised of two components.

- **Radio Sensor**: Receives and optionally sends radio signals. Computes some meaningful Radio Sensor measurements based on the received signals.
- **Inference System**: Analyzes the Radio Sensor measurements to derive the context of interest. Examples of inferred contexts using the Radio Sensor include subject location [45], breathing rate [31], gestures [81], among others. In our case the context is the activity of a subject.

The Radio Sensor is deployed in the environment. In this environment a subject conducts an activity, which the Inference System should recognize. Figure 7a shows a schematic illustration of such a system installation. Here, the Radio Sensor consists of four transceivers (circles) which are spread across the room. Measurements by the sensor are passed on to the Inference System which performs the analysis and recognizes the activity.

Figure 7b provides a more abstract view on the nature of a measurement. It emphasizes different aspects which are likely to influence the measurement. Here,




(a) Implementation Perspective (b) Abstract Perspective

Figure 7.: Perspectives on a device-free, radio-based Activity Recognition System

a measurement represents the data provided by a Radio Sensor(RF) biased by the parameter configuration of this sensor biased by the parameters of the environment(Env) biased by a subject(Subj) which can be described by a number of parameters and which interprets an activity(Act) at some point in time $t$:

$$\text{Measurement}(t) = \text{RF}_{\text{param}(t)}(\text{Env}_{\text{param}(t)}(\text{Subj}_{\text{param}(t)}(\text{Act}))) \tag{5}$$

For better understanding the idea of these influences, consider the following settings for parameters of each aspect: The Radio Sensor may be using either two transceivers or twenty. The environment may be a small office room or a large lecture hall. The activity may be performed by a child or a grown up. Each combination of these parameters will likely lead to another outcome in the measurement and consequently in the recognition. Of course, the configuration of the Inference System will also strongly affect the actual recognition performance. But to increase the chance of Activity Recognition, it is crucial that the raw data provided to the Inference System contains as much information about the actual activity as possible. The underlying assumption of this chapter is that this information content is influenced by parameters related to the aforementioned aspects. Hence, in the investigations in this chapter we exclude the optimization of the Inference System component. This is explored starting from Chapter 5. Note that in the described perspective we consider an activity not as an aspect but as an atomic information. Other information which may typically be seen as attached to the

activity such as the execution speed or the actual performance are parameters we consider attached to the subject. Hence, in the optimal case the Inference System can provide the exact *Activity* originally embedded in the measurement.

Having reviewed the related work and having discussed these aspects in more detail, we selected five parameters for further investigation. For each parameter setting we employed the full Activity Recognition tool chain to judge their impact using a total of 54 evaluation data sets with a single link setup and three different activities. Using a newly developed criteria based on accuracy range the different influences of the parameters were demonstrated. Selected parameters and their influence were as follows: size of environment (range: 17%), transmission power (16%), frequency spacing (15%), frequency diversity/number of frequencies (14%) and length of wireless link (11%). For parameters which cannot be adapted after the system installation the results provide a best practice guideline. For parameters which can be tuned during runtime, this may allow to increase the information content of activities in the measurement and thus, improve recognition performance. The definition of aspects of influence and attached parameters opens a door to a more structured investigation methodology for device-free, radio-based recognition.

## 4.2 RELATED WORK

Literature on parameters influencing Activity Recognition performance is scarce. However, other radio-based research in domains such as localization or motion detection conducted studies which may provide useful insights and are therefore included here. Parts of this overview were published in Ref. [75].

### 4.2.1 *Subject Location*

It has been shown in a large corpus of research that device-free localization of subjects is feasible with very good accuracy, e.g. [67, 11]. However, in this section we review work which describes the actual effect of a subject's location in relation to the position of the transceivers. Zhang et al. [101] installed a number of radio modules in a grid in 2.4 m height on the ceiling. They observed that the area of influence of a wireless link is an elliptical area on the floor. They reported that the strongest impact of human motion is located in the midpoint of this ellipse. Yao et al. [98] described an inverse relationship of distance between radio transceiver and subject and the observed RSS variance. The contour plot created based on their experiments showed that the strongest effect on RSS was close to the transceivers or along the line of sight (LoS). In Ref. [82] we employed Software Defined Radios in a distance of 2 m for activity recognition. We conducted four activities (crawling, walking, lying and standing) in 11 locations around the receiver with only some locations in line of sight (LoS). Best performance was reached for activities which were close to the receiver or close to the LoS. In Ref. [59] Wilson and Patwari present a model predicting RSS variance of a moving subject in

relation to transmitter and receiver position and installation height. According to the model, variance is high when a subject moves in the LoS in proximity to either of the modules. Subject influence fades with increasing distance from the nodes. Hence, the model fits to the observations of Ref. [82] and Ref. [98] given the correct parameterization. The authors claimed [59] that it can also explain the observations in Ref. [101] although subjects were stationary in the latter experiments.

### 4.2.2 *Carrier Frequency*

Device-free, radio-based context recognition research has employed various carrier frequencies successfully [75]. However, only few works have reported on comparisons between carrier frequencies. Reschke et al. [64] used Software Defined Radios on 900 MHz and 2.4 GHz for activity and situation recognition. Results were improved when using a frequency of 2.4 GHz. Woyach et al. [93] compared how RSS is affected when a human moves in the LoS with respect to two different carrier frequencies. They reported that the fluctuation of 2.4 GHz signals was about twice as high as the fluctuation of 433 MHz signals. In Ref. [41], Kaltiokallio et al. compared the performance of their localization system in respect to the used IEEE 802.15.4 2.4 GHz frequency channel. They showed that localization accuracy could be considerable increased when choosing an appropriate channel.

### 4.2.3 *Link Length and Transmission Power*

Zhang et al. further conducted experiments investigating subject influence in respect to radio transmission power and link length [101, 102]. They observed that for relatively short distances (2 m, 3 m) using a lower transmission power results in an increased impact of a stationary subject in LoS. However, for longer distances (4 m, 5 m) between radio modules they did not observe a clear difference in induced signal fluctuation in relation to transmission power. In general, their experiments showed a stronger effect of a subject for the 2 m or 3 m links than for the 4 m and 5 m links.

### 4.2.4 *Density and Topology*

Kosba et al. investigated the impact of three different transceiver layouts (topologies) on the accuracy of device-free localization [46]. In these investigations the number of transceivers (density) was constant but the spatial arrangement was changed. The authors demonstrated a significant effect of the topology on the localization error.

Patwari et al. demonstrated that RSS-based breathing detection could be improved when the number of radio modules was increased [60]. They also showed that using directional antennas improved the detection accuracy. Similarly, Reschke et al. found that in their Activity Recognition experiments accuracy was improved when employing three instead of two SDR [64].

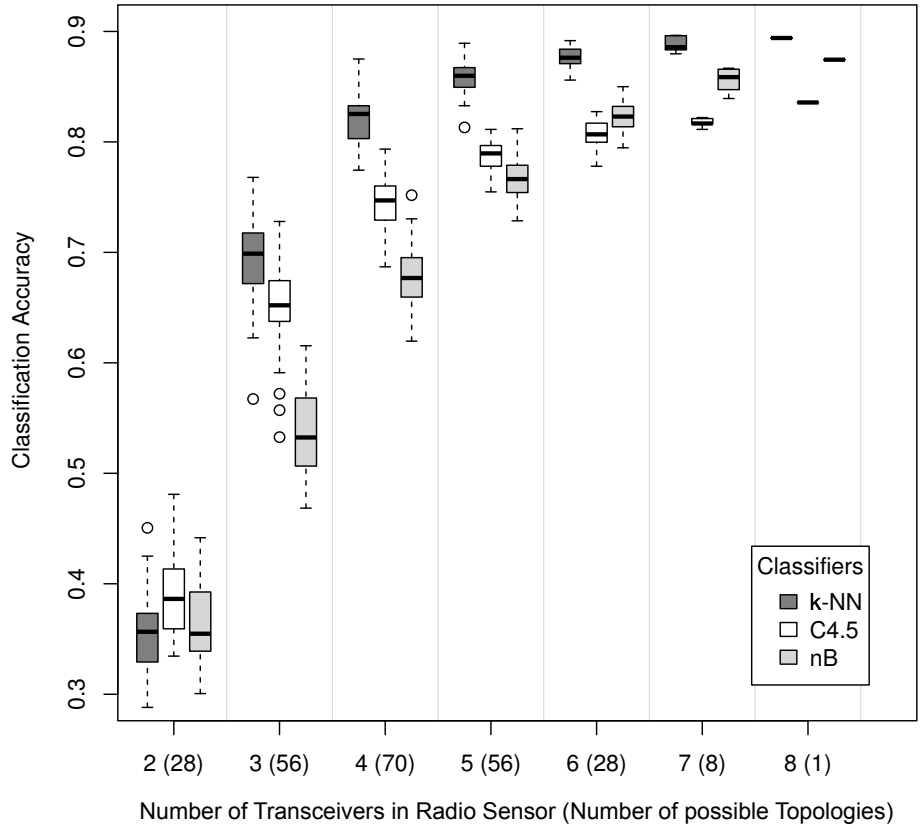



Figure 8.: Radio Sensor density and topology vs recognition accuracy from Scholz et al. [73]

In our paper [73], we investigated the influence of density and topology for device-free Activity Recognition. The test bed was a medium sized office room with a total of 8 transceivers installed. Therein 7 activities were conducted. The data was analyzed using cross-validation and three typical classifiers (naive Bayes: nB, decision tree C4.5 and k-Nearest Neighbors: k-NN). Each classifier was tested on each possible transceiver combination (density and topology). Figure 8 shows a box plot where each box represents all possible topologies for a given number of transceivers. When choosing the correct four transceivers a performance comparable to the overall best performance with 8 transceivers can be achieved (cf. Figure 8 upper whisker of k-NN for 4 nodes vs 8 nodes). Hence, RSS-based Activity Recognition accuracy generally improves with increasing transceiver density. In addition, depending on the other aspects specific topologies exist which allow comparable recognition rates even with lower density.

## 4.3 ANALYSIS

In this section we list typical parameters for each of the mentioned aspects. The goal was to select a number of parameters for the investigation and to develop an evaluation approach and metrics to judge their impact.

### 4.3.1 *Description of Parameters of the identified Aspects*

In the introduction we have identified three aspects. In this section we describe these aspects more closely and give an overview over associated parameters. Note that for all aspects parameter values may change over time.

*Subject*

The subject aspect refers to the subject performing an activity. While the activity would not exist without the subject conducting it, the subject is also the first level of distorting the information of the activity in the radio signal. For instance, through subtle changes in each activity repetition. The following list presents a sample of the parameters associated with the subject aspect.

- Location: The position of the subject in the environment.
- Clothing: The clothing of the subject.
- Activity Execution: A meta parameter related to the execution of the activity by the subject. This may be approximated by speed or orientation of the subject. Ideally the data includes the actual motion vector and trajectories, i.e. postures, the subject performs while conducting the activity.
- Dimension: A meta parameter describing the subject's spatial expansion as well as the volume and materials of the subject's tissues. These may be approximated using height, weight, width, sex, morphological and biometric parameters, among others.

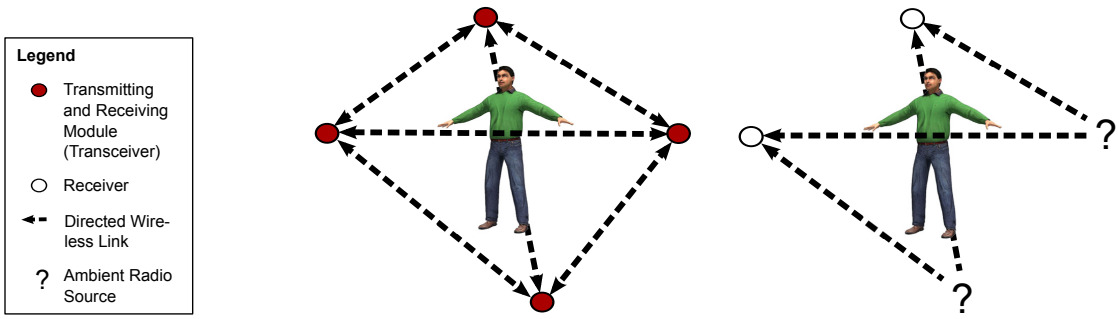
*Environment*

The environment describes the surrounding space in which the Radio Sensor is installed. The following list presents a collection of parameters associated with this aspect.

- Dimension and Material: Meta parameter referring to aspects such as width, height, depth and material of the surrounding environment. The surrounding environment comprises all furniture, rooms or other structures which are within a relevant distance of the Radio Sensor, i.e. within a distance in which they affect the sensor measurements. In other words, if the sensor values do not change when a specific structure is removed then this part of the environment does not have to be included in this description. The only exception is the subject performing the activities. However, note that additional subjects in the relevant area are considered parts of the environment as we focus on single user activity recognition in this work.
- Climate: Meta parameter referring to climate related attributes which may change Radio Sensor readings. Examples include temperature, humidity and wind.

*Radio Sensor*

The Radio Sensor is a configuration of a single or multiple hardware entities (and firmware) receiving radio signals. We distinguish two general types of Radio Sensor [75, 82]: A Radio Sensor actively transmitting radio signals is called *active* (Fig. 9a). A Radio Sensor only analyzing ambient radio signals e.g. from FM radio stations is called *passive* since it is not actively transmitting (Fig. 9b).




(a) Active Radio Sensor (b) Passive Radio Sensor

Figure 9.: Radio Sensor Configurations

Clearly, a system which is passive may be considered calm or less obtrusive in comparison to an active system as it does not emit electromagnetic radiation. In contrast, we consider active systems more stable and robust as the senders are all part and under control of the Radio Sensor. I.e. as long as the sensor is operational, senders are available. This cannot be assumed for passive sensors. For instance, an FM radio station or WiFi access point might shut down or change their transmission frequency or some other important parameter. In addition, ambient radio sources may be located far away from the actual monitored area, leading to signal changes not associated with the context of interest but with external influences. Therefore, we further consider the passive Radio Sensor more noisy. Further important Radio Sensor parameters are described in the following list.

- Type: Passive or active Radio Sensor as described above.
- Transceiver: A meta parameter which is the actual name and identification information of the used hardware implementation. It refers to many more parameters such as the receiver sensitivity, bandwidth, measurement resolution, radio frequency hardware design, among others.
- Antenna: Antenna specification/characterization and orientation.
- Measurement: The quantity which is measured. Examples include received signal strength (RSS), time of flight, angle of arrival, among others.
- Carrier Frequencies: This parameter describes the carrier frequencies on which the sensor conducts measurements.
- Transmission Power: The power which is used in the transmission (amplitude of the radiated electromagnetic waves).

- Sampling Rate: Frequency of sensor measurements.
- Sampling Scheme: Describes the used measurement scheme, i.e. which receiver receives information from which sender(s) at which time.
- Density and Topology: Meta parameter which encodes the spatial configuration of transceivers. More specifically it includes the number of transceivers, their position and orientation. Based on these information it keys further information such as the length of LoS links and heights above ground of the transceivers.

### 4.3.2 *Selection of Parameters for Investigation*

In the previous section we have explored parameters associated with the three aspects of a device-free, radio-based Activity Recognition system. However, capturing all these parameters would be to complex and thus, is not feasible. Previous research has shown that recognition systems can be successfully realized by implicitly including parameters, typically through the means of calibration. Hence, we focus on parameters that are assumed to have strong effects on the recognition, which can be controlled (e.g. to keep them constant during experiments) and can be acquired with reasonable effort.

#### *Radio Sensor*

The Radio Sensor is the aspect over which we exhibit most control. Even after deployment in an unknown environment some parameters of the sensor are likely to be known and could probably be adapted to suit the environment (e.g. transmission power). Hence, from the perspective of a radio-based recognition system investigating Radio Sensor parameters seems most beneficial.

As previously described two types of Radio Sensors are possible. In passive operation the sensor is calm but offers less control and introduces additional uncertainties. For this reason we only considered active sensors.

As described in the related work section, both density and topology were investigated previously [73] and findings therein correspond to reports of other research and intuition: Increasing density increases coverage and in conclusion typically allows for better recognition rates. Alternatively, the topology may be optimized for the specific activities leading to similar recognition rates.

Ref. [60] showed that radio-based recognition can benefit from directional antennas. This seems intuitive as the influence from external noise is reduced and changes in radio signals originating from the area of influence are emphasized. As they have already been explored, we exclude antenna and density from the parameter investigation and consider the following Radio Sensor parameters for evaluation:

- Link length

- Transmission power
- Frequency diversity
- Frequency spacing

Employed and tested parameter settings were derived from typical indoor environments (link length), transceiver options (power, diversity, spacing) and availability in Smart Home technologies.

*Environment*

The environment is an aspect over which the recognition system has no control and possibly zero knowledge. In addition, parts of the environment change continuously (e.g. door). For this reason, we consider investigations related to environmental parameters as explorations of device-free Activity Recognition limitations and best practices. One of the environmental parameters which is suggested to have a significant impact on recognition is the size. Therefore we consider this as investigation parameter.

- Size of the environment

*Subject*

Like the environment, the subject is an aspect over which the recognition system has no control and typically zero knowledge. Some parameters of the subject undergo continuous change. Frequent parameter changes are typically induced by the execution of physical activities (e.g. walking affects location). Such parameter changes possibly influence signal propagation leading to recognizable patterns in the measurement. Hence, variations of frequently changing subject parameters are inherent to the parameter investigation as different activities and locations are tested. For the evaluation we defined the following parameter settings:

- Activities: Walking, Standing, Standing and Waving
- Locations: 3 different locations

The three activities were selected as representatives of each of the identified activity categories (cf. Chapter 8). In contrast to parameters affected by physical activities, other parameters such as height or width are rarely influenced and hence change less frequently. Thus, these may be considered subject specific parameters. We investigate such parameters in order to differentiate between subject classes in Chapter 7.

### 4.3.3 *Evaluation Approach*

For evaluation a list of default parameter values was defined. Default values are settings which are kept constant while the investigated parameter is changed

and the activities defined in the previous section are conducted. The current parameter setting was evaluated using an Inference System based on the standard Activity Recognition process [8] and appropriate metrics. To enhance significance of the evaluation independent training and test sets were used [10]. Hence, we recorded multiple independent annotated data sets which contained Radio Sensor measurements of the subject conducting the defined activities. This data was then used to train a classifier which was then evaluated on the other data sets. Then this process was repeated with the remaining data sets. In order to compensate for possible outliers during the experiments, but keep the evaluation feasible we chose to record three data sets per parameter setting.

*Inference System Configuration*

The focus of this chapter lies on parameter identification and evaluation of the Radio Sensor and the environment aspect. For this reason the configuration of the Inference System was chosen based on best practices. An SVM classifier was reported to improve recognition for Activity Recognition in SDR experiments [78] over other classifiers such as C4.5. However, as we show in Chapter 8 the classifier is much less important than the feature. As features we chose standard deviation as it was reported to be a robust feature for motion detection [45] and delivered promising results in our cross validation investigation [73]. The window settings are also from Ref. [73]. Thus, we use the following settings for the Inference System:

- Classifier: Support Vector Machine (SVM)
- SVM kernel: Radial Basis Function
- Feature: Standard Deviation
- Feature window size: 40 samples (1 s)
- Feature window overlap: None

*Metrics*

The accuracy is an overall metric to assess the performance of a trained classifier (cf. Eq. 3). Since we wanted to evaluate multiple data sets for the same parameter setting, we compute the average accuracy given a single training set $r$ vs the remaining test sets (AvgAcc, Eq. 6). In this equation, $|\,\text{Data Sets}\backslash r\,|$ refers to the number of elements in Data Sets excluding the training data set $r$. The variable $s$ denotes the current data set under test.

$$\text{AvgAcc}(r) = \frac{\sum_{s}^{\text{Data Sets}\backslash r} \text{Accuracy}(r,s)}{|\,\text{Data Sets}\backslash r\,|} \tag{6}$$

In order to assess the performance of a specific parameter setting across all train/test combinations we compute the average over the averaged accuracies (AvgAvgAcc, Eq. 7) and the average over the standard deviation for each combination (AvgStdAcc, Eq. 8).

$$\text{AvgAvgAcc} = \frac{1}{|\,\text{Data Sets}\,|} \sum_{r}^{\text{Data Sets}} \text{AvgAcc}(r) \tag{7}$$

$$\text{AvgStdAcc} = \frac{1}{|\,\text{Data Sets}\,|} \sum_{r}^{\text{Data Sets}} \sqrt{\frac{\sum_{s}^{\text{Data Sets}\setminus r} \left(\text{Accuracy}(r,s) - \text{AvgAcc}(r)\right)^2}{|\,\text{Data Sets}\setminus r\,|}} \tag{8}$$

The ultimate goal of this chapter is not only to evaluate the effect of different settings of a single parameter but also to compare different parameters to estimate their importance. Based on the rationale that a parameter should be more carefully chosen if it can inflict a strong decrease in accuracy, we define parameter importance as the observed accuracy range (AccRng, Eq. 9) over the tested settings of this parameter.

$$\text{AccRng} = \underset{\text{parameter setting}}{\text{argmax}} \left(\text{AvgAvgAcc}\right) - \underset{\text{parameter setting}}{\text{argmin}} \left(\text{AvgAvgAcc}\right) \tag{9}$$

## 4.4 EXPERIMENTAL EVALUATION

In this section we define and summarize the default parameters and the experimental setup. We formulate an expectation about the parameter's influence and investigate each parameter experimentally. Finally, we employ the AccRng metric to determine the impact of each parameter.

### 4.4.1 *Default Parameters*

Apart from the investigated parameter, all other parameters are fixed to default values (Tab. 3).

Experiments are conducted using a standard setup (Fig. 10). All activities were performed on either one of the three positions or evenly spread across all three locations (walking). To force the system to distinguish between the activities and not the locations, the three locations of the two in-place activities were not labelled in the training and test data sets.

| **Aspect**/Parameters | **Default Values** |
|---|---|
| Activities | Standing, Waving, Walking |
| **Subject** | |
| Activity Duration | 30 s (in-place activities: 10 s per position) |
| Activity Speed | Approx. constant |
| Sex, Height, Weight | Male, 1.83 m, 81 kg |
| Location | Three different positions (Fig. 10) |
| Orientation | Always the same direction, shoulders orthogonal to LoS (Fig. 10) |
| **Environment** | |
| Size | Medium room (23 $m^2$x3.30 m height) |
| Material | Brick |
| Furnishing | None |
| **Radio Sensor** | |
| Type | Active |
| Link Length | 3 m |
| Installation Height | 1.25 m |
| Sampling Frequency | 40 Hz |
| Topology | Single Link (Two radio modules) |
| Carrier Frequencies | 2326 MHz, 2425 MHz, 2524 MHz |
| Carrier Frequency Spacing | 100 MHz |
| Transmission Power | +4 dBm |
| Antenna | Upright omnidirectional $\lambda/2$ dipole antenna |
| Measurement | Received Signal Strength (RSS) |
| Transceiver | AT86RF233 |
| **Inference System** | |
| Classifier | Support Vector Machine (SVM) |
| SVM Kernel | Radial Basis Function |
| Feature | Standard Deviation |
| Feature Window Size | 40 samples (1 s) |
| Feature Window Overlap | 0 samples |

Table 3.: Default Settings for the Parameter Investigation Experiments

### 4.4.2 *Link Length*

In this section, the distance between the devices and its effect on the recognition accuracy is examined.

- **Background and Expectation:** As described in Chapter 2, during propagation the electromagnetic energy dissipates proportional to the square of the distance and frequency (Eq. 1). In other words, if the distance between two transceivers is doubled signal path loss will quadruple. The reason for this effect is that the surface of the propagation sphere on which the emitted energy is distributed increases fourfold for each increase in distance. Hence, the energy received by an antenna in this distance will be decreased accordingly. Now imagine that instead of an antenna, a subject is introduced at a

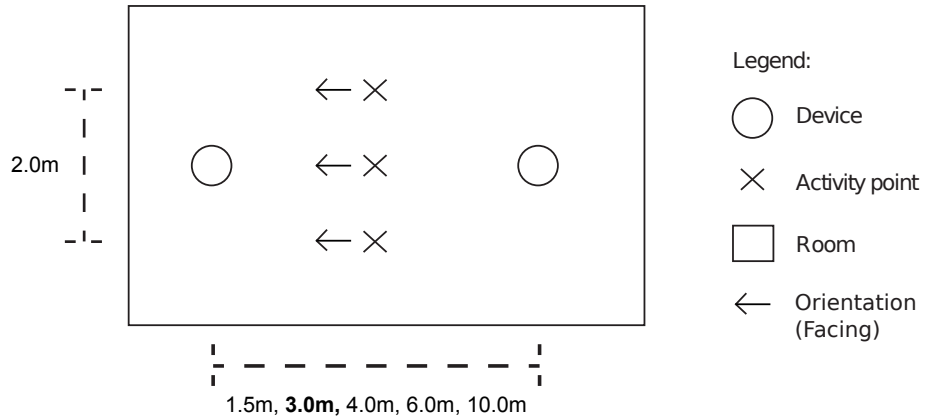



Figure 10.: Setup for parameter experiments (default link length is 3.0 m)

distance $d$ from the sender. Depending on $d$ the subject will affect a larger or smaller portion of the propagation sphere. Thus, there is a direct relation of the impact of a subject on the signal and its distance to the transceivers. For a single link as used in the evaluation the subject's least impact on the line of sight (LoS) must be in the middle as the distance in LoS from both sender and receiver is maximal. Moving the subject closer to either transceiver will increase the impact. In our experiments, activities were conducted in the middle between transmitter and receiver. Thus, the distance between subject and transceivers increases with $\frac{d}{2}$ as the distance between the radio modules increases. Therefore we assume that recognition performance decreases with increasing distance.

- **Experiment and Results:** Using the previously defined default parameters we performed and recorded activities for a total of 5 different distances. The tested distances were 1.5m, 3.0m, 4.0m, 6.0m, 10.0m. For each distance we conducted three recordings of the specified activities. The standard setup of the experiment (cf. Figure 10) was adapted for different link lengths such that the activity points were always in the middle of LoS. At the same time, the outer activity points were left at 1 m distance from LoS. To conduct all investigations in the same space we used a different room than specified in the default parameters. The area of the room was 72 $m^2$ (6x12 m). Thus, in all experiments objects were at least 1 m away from the radio modules.

  Figure 11 shows the results of the experiments. We observe that the average accuracy increased steadily until reaching 90% at 4.0 m. Thereafter accuracy decreases (6.0 m) before it recovers slightly (10.0 m). For the tested settings the optimal distance is 4.0 m having the highest average accuracy and a small standard deviation. For the tested distances minimum average accuracy was 79% and maximal achieved accuracy was 90%. The observed range (AccRng) over the average accuracies induced by link length was 10.8% points.

  This is opposed to the expected downward trend in accuracy from short to long distances. From observations in the raw data we find that at the

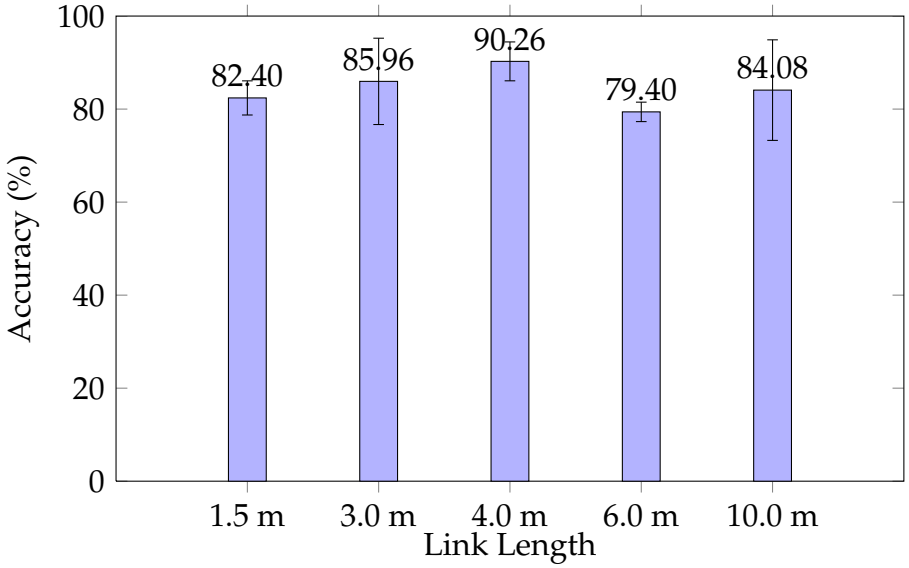



Figure 11.: Accuracy (AvgAvgAcc) vs device distance. Error bars indicate standard deviation (AvgStdAcc).

shortest distance (1.5 m) even small differences in activity execution induce changes in the RSS. While this may allow to distinguish the different ways an activity was conducted it impairs recognition when the same activity is executed slightly differently. For links >1.5 m, these intra-activity differences are increasingly blurred improving recognition performance (3 m, 4 m). At 4.0 m the intra-activity difference is minimal while activity classes still leave a characteristic footprint in the RSS, leading to best recognition performance. Increasing the distance further (6 m) distorts this characteristic footprint, explaining the drop in recognition performance. However, against our expectation performance is improved in the 10m experiment. Here, we hypothesize that the proximity to the brick walling of the room (room length: 12.0 m) added further propagation paths which increased system sensitivity.

- **Conclusion:** The influence of link length is non-trivial and associated with the surrounding environment. In general, longer links and very short links will reduce recognition accuracy.

### 4.4.3 *Radio Transmission Power*

- **Background and Expectation:** The received power falls in relation to frequency and distance. Increasing the power will obviously not change this general behavior. However, increasing power will increase the relevance of multipaths and therefore increase the (measurable) radio signal coverage of the area. Relevance of multipaths means that paths which were below the sensitivity of the receiver may now become important enough to have an effect on the received signal. In other words, in low power mode only paths

which have similar length like the shortest path (LoS) are important. When power is increased also propagation paths which diverge from LoS and e.g. paths with reflections become more important. Thus, space between the transceivers is more strongly illuminated by electromagnetic radiation. For this reason we assume that recognition accuracy improves with increasing transmission power.

- **Experiment and Results** Using the aforementioned setup we performed and recorded activities using three different power settings. The power settings were evenly spread in three steps across the range of available settings from -17 dBm to +4 dBm. For each setting three activity data sets were recorded. Figure 12 shows the results of the evaluation.

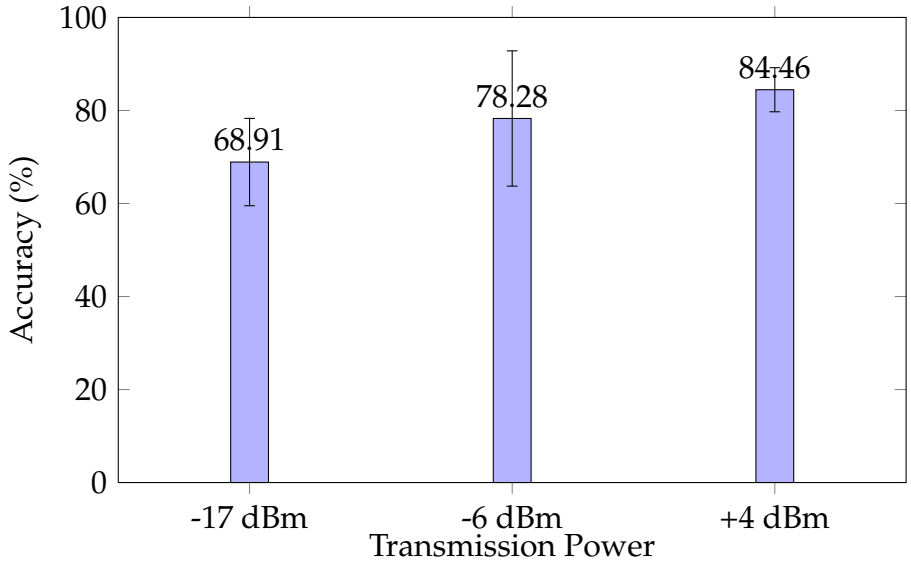



Figure 12.: Accuracy (AvgAvgAcc) vs transmission power. Error bars indicate standard deviation (AvgStdAcc).

Recognition accuracy increases with transmission power while variability of accuracy decreases. The lowest power settings achieved an average accuracy of 68.91%, the highest power setting achieved an average accuracy of 84.46%. Accuracy range inflicted by varying the power setting by 21 dBm is 15.55% points. In the experiments, an increase of approx. 10dBm leads to an improvement of 6-10% accuracy depending on the original transmission power. Accuracy gain is higher when the original transmission power is lower. Note that an increase of 10dBm means that the actual power is increased by a factor of 10. Thus, considering the absolute power the accuracy gain from -17 dBm (0.02 mW) to -6 dBm (0.25 mW) is more efficient (10%) than the gain from -6 dBm to 4 dBm (2.50 mW), which is only 6%. Hence, the returned accuracy gain decreases with increasing power. This observation may indicate an increasing saturation of the test bed with electromagnetic radiation.

- **Conclusion:** Increased power increases coverage by making alternative radio paths more influential in the RSS. Hence, spatial changes in the illuminated area will lead to stronger changes in the received power. Therefore the sensitivity for activities in the Radio Sensor increases.

### 4.4.4 *Frequency Diversity*

In this section we report from the influence of increasing the number of measurement frequencies for the same physical link.

- **Background and Expectation:** Literature suggested that frequency diversity provides additional information about the area of interest [41]. The reasons for the additional information lie in the different propagation of waves at different frequencies. Generally, propagation of electromagnetic waves at different frequencies, i.e. waves with different wavelengths, lead to different interference patterns. In addition, electromagnetic waves also interact differently with objects in the propagation paths with respect to the carrier frequency. For instance, the angle of refraction depends on the wavelength of the incident wave [63]. Thus, as an extreme case consider a propagation path for one frequency which strikes the area of a subject while on another frequency it does not. Additionally, the antenna gain typically vary across frequencies. For this reason, our expectation is that measuring on multiple frequencies increases the recognition accuracy.

- **Experiment and Results:** We recorded a total of nine data sets. Each set was recorded with all three frequencies. The frequencies were: $f_1$ = 2326 MHz, $f_2$ = 2425 MHz, $f_3$ = 2525 MHz. We used the first three data sets to evaluate the accuracy using a single frequency. Hence, we computed the average over the accuracy when consecutively training and testing with any one of the frequencies, i.e. $\frac{\sum \text{AvgAvgAcc}(f_i)}{3}$ and similarly for the standard deviation. We used the next three data sets to evaluate the accuracy using two frequencies. Hence, we computed the average over the achieved accuracy when the system is trained/tested using any combination of two out of the three frequencies as features, i.e. $\frac{\text{AvgAvgAcc}(f_1,f_2)+\text{AvgAvgAcc}(f_2,f_3)+\text{AvgAvgAcc}(f_1,f_3)}{3}$, similarly for AvgStdAcc. Finally, we compute the accuracy for the system when employing all three frequencies. Note that the averaging is performed after the system was evaluated using the hold-out evaluation approach described in the analysis section. Figure 13 shows the result of the evaluation.

  We observe that an increase in the number of measured carrier frequencies improves recognition accuracy. However, while the change from a single frequency measurement to two frequencies improves recognition by 8%, the improvement to three frequencies provides only a 6% points increase. The observed average accuracy range over all tested parameters settings was 13.79% points.

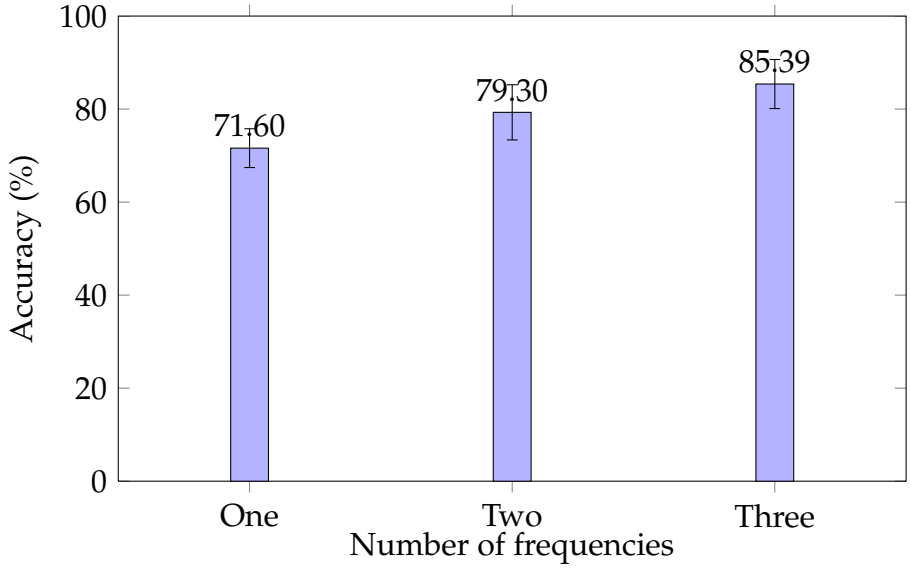



Figure 13.: Accuracy (AvgAvgAcc) vs frequency diversity. Error bars indicate standard deviation (AvgStdAcc).

- **Conclusion:** In accordance with our expectation we find that accuracy increases when the number of measured frequencies is increased. This approach may have further potential to increase accuracy.

### 4.4.5 *Frequency Spacing*

- **Background and Expectation:** The previous experiment has shown that frequency diversity strongly improves recognition. Following the previous assumption we postulate that the more carrier frequencies differ the more likely are the propagating waves to be influenced by different characteristics of an activity. Hence, the system will be provided with more information when frequencies are spaced further apart and recognition performance will improve.

- **Experiment and Results:** The radio hardware used for the experiments has a bandwidth of 203 MHz (from 2324 MHz to 2527 MHz). In respect to the default settings the Radio Sensor operated on three carrier frequencies. Based on the provided bandwidth we tested the following three frequency spacings:

  - Low frequency spacing (25 MHz): 2475 MHz, 2500 MHz, 2525 MHz
  - Medium frequency spacing (50 MHz): 2425 MHz, 2475 MHz, 2525 MHz
  - Maximal frequency spacing (100 MHz): 2325 MHz, 2425 MHz, 2525 MHz

  For each spacing we recorded three data sets which were evaluated using the previously described procedure. The results are shown in Figure 14.

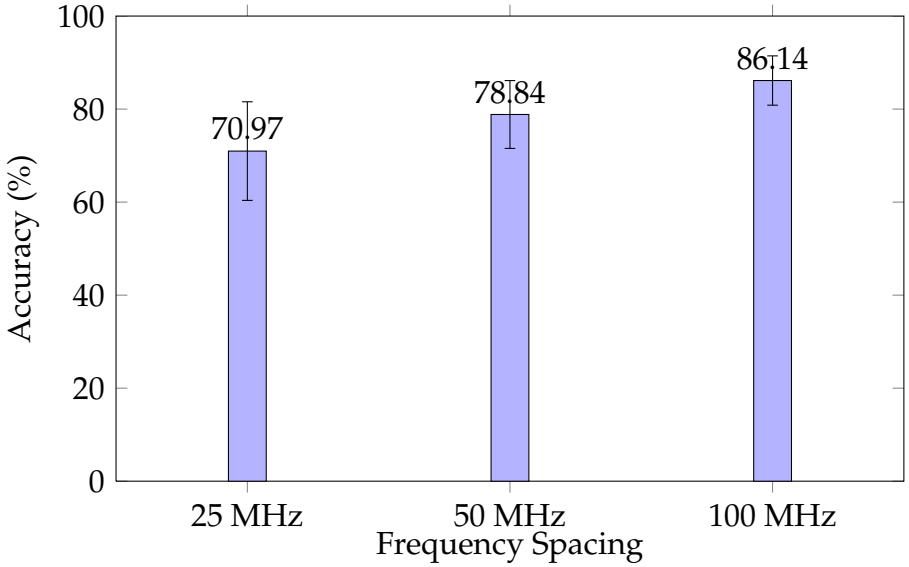



Figure 14.: Accuracy (AvgAvgAcc) vs frequency spacing. Error bars indicate standard deviation (AvgStdAcc).

For the lowest frequency spacing of 25 MHz we achieve an average of 70% accuracy. 50 MHz spacing increases accuracy by 8% points. By doubling the spacing again to 100 MHz we gain another 8% points on accuracy. The difference between the lowest and best achieved average accuracy is 15.17% points.

- **Conclusion:** Increasing frequency spacing improves recognition performance.

#### 4.4.6 *Size of Environment*

- **Background and Expectation:** In free space the energy emanating from the antenna of a sender freely propagates in all directions. In this case only few electromagnetic waves propagating in the line of sight impinge on the receiving antenna. Due to propagation effects such as reflection, transmission and diffraction the introduction of new objects (floor, walls, furniture) to the surroundings increases the likelihood of additional paths reaching the receiver. These additional paths also introduce new trajectories through space possibly crossing a subject's position. Hence, recognition in an indoor scenario should generally be better than in an outside scenario. Additionally, the distance of these new objects from the transceivers is important. Reflected waves travel from the sender to the object and then back to the receiver. Hence, the closer the objects to the setup the more likely that the newly added path has a large enough amplitude to have a significant effect in the received energy. The further away from the setup the object is located the less likely will the reflected wave yield a great difference in the received

energy. Thus, we expect that recognition is less accurate in larger rooms than in smaller rooms.

- **Experiment and Results:** The experiment was conducted in four environments of different size. For all environments except the smallest room, no obstacles were within an area of at least 1 m around the radio modules and the activity area. In the smallest room only the sender was 1 m away from the wall, while the receiver was much closer to the wall. However, the room has an inlet of approx. 1 m at the height of the receiver. The smallest room was also too narrow to stand at exactly 1 m distance from LoS. Hence, the subject stood at the closest position next to this point. In each environment three data sets were recorded an evaluated using the described procedure. Figure 15 shows the results of the evaluation. The best average accuracy of

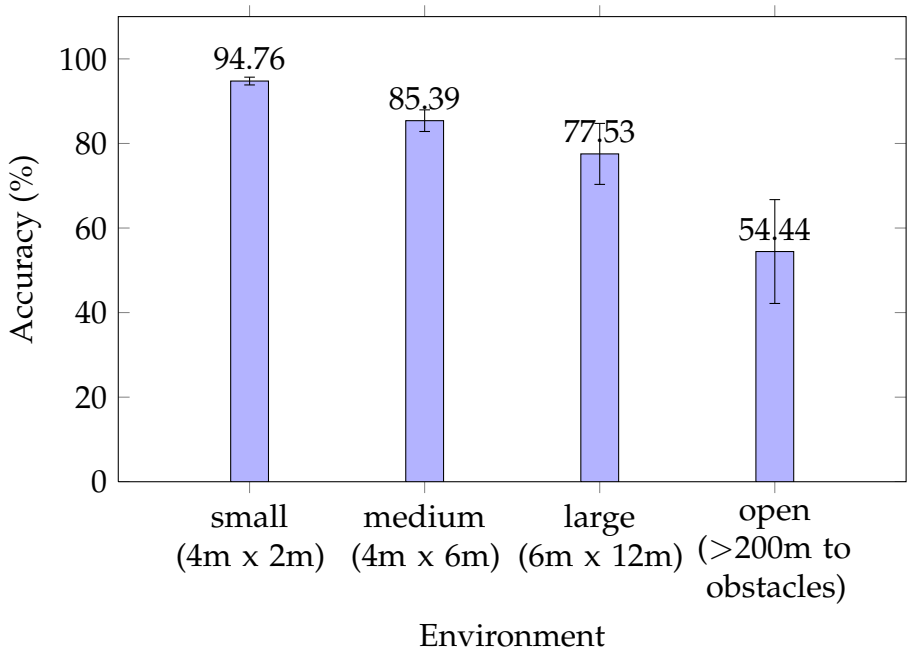



Figure 15.: Accuracy (AvgAvgAcc) vs size of environment. Error bars indicate standard deviation (AvgStdAcc).

94.76% and lowest standard deviation were achieved in the smallest room. The lowest accuracy of 54.44% and highest standard deviation was observed in the open environment. Note that the accuracy achieved for the smallest room is probably also partly due to the subject standing closer to line of sight than in the other settings.

The determined accuracy range over all tested environments is 40.30% points. If only considering the indoor environments the range is 17.23% points.

- **Conclusion:** We find our expectation fulfilled observing that with increasing room size the accuracy is reduced while standard deviation is increased.

### 4.4.7 *Comparison of Parameters*

In this section we compare the evaluated parameters with respect to their importance. As described in the analysis we use the range over the accuracies for this purpose (AccRng). Figure 16 shows the accuracy range for each parameter. Note that we have excluded the open environment from the parameter "Size of Environment" as this seems irrelevant for device-free Activity Recognition systems in the Smart Home. All parameters are relatively close with only a difference of 7% points from the parameter with the lowest impact (Link Length) to the most influential (Size of Environment). Albeit excluding the open environment from the comparison "Size of Environment" is the most influential parameter with a difference of 2% from the next parameter. Since it is a parameter of the environment aspect it cannot be controlled. However, the evaluation results may serve as base for best practice considerations, i.e. to estimate what to expect of an installed system in an area of certain size. For medium-sized indoor environments (>4x6 m) it is advised to utilize multiple links. On the other hand, in smaller room a single link could suffice.

All remaining parameters are related to the Radio Sensor aspect. Transmission power is the parameter with the second largest influence. Therefore the default for a device-free Activity Recognition system should be the use of the maximal possible transmission power. Frequency spacing is the second most influential parameter of the Radio Sensor. As for the transmission power maximal settings should be used, i.e. maximizing spacing between frequencies as this promises best results. Similarly, frequency diversity has a reduced but still important influence. Thus, for an Activity Recognition system it should provide the option for measuring at as many frequencies as possible and these should be maximally spaced over the available bandwidth. Lastly, the link length parameter is one of the few parameters of the Radio Sensor not easily adaptable when following the vision of implementing the sensor on top of a Smart Home infrastructure. A best practice consideration for this parameter is to avoid links longer than 4m as these are often less sensitive to subject activities.

## 4.5 DISCUSSION

### 4.5.1 *Aspects*

In this chapter we defined three primary aspects which influence a measurement. These aspects are ordered in a layered hierarchy in which outer layers may change the influence of an inner layer. This definition is clearly not perfect. For instance, it does currently not include the case of multiple subjects. However, while parameter investigations were conducted previously [73], it is now possible to relate them to a specific aspect; enabling a structured investigation methodology. Such a structured investigation can have multiple goals. Two of these are presented in this chapter: deriving best practices and ranking parameters for their importance.

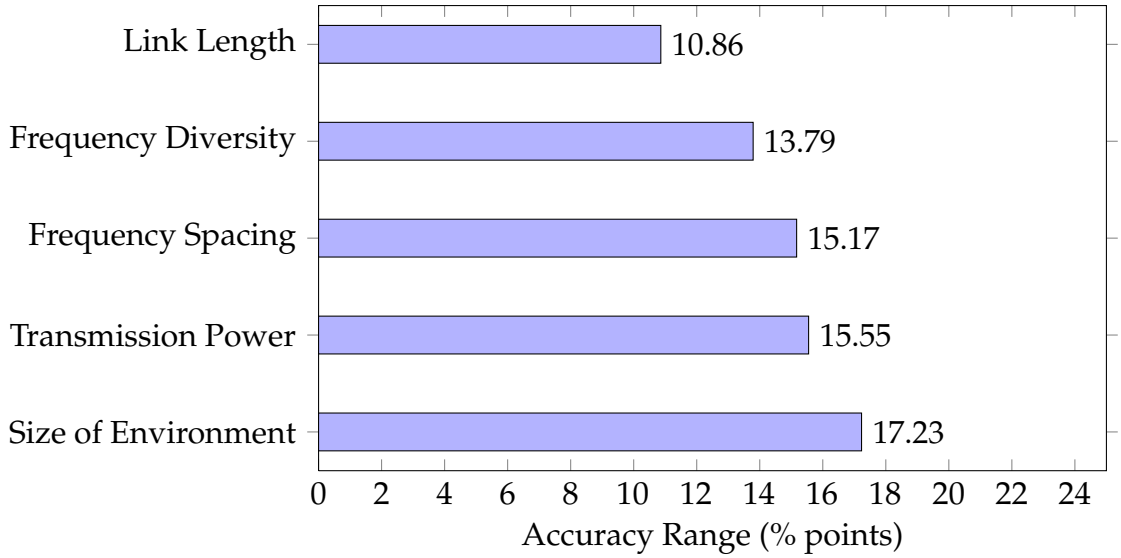



Figure 16.: Accuracy Range (AccRng) for each investigated Parameter

For instance, to focus system development on more important parameters. The ultimate goal is to develop a model which considers the interplay of all parameters. This would allow to build parameter-based recognition algorithms and to select optimal parameters for specific recognition challenges.

### 4.5.2 *Experiments*

The experiment setup was very simple considering only a selection of three activities and a single link. Nevertheless, we believe that even with variations of in-place activities the tendency in the results would be similar as most parameters induced some change in the electromagnetic coverage of the room. For some of the parameters a larger number of settings would have been beneficial, e.g. to determine the optimal setting. However, optimal settings are also a function of the different aspects and might be restricted by the used parameter set. A number of parameters was not investigated, such as the installation height of the transceivers or subject specific characteristics. However, for the installation height in respect to activity or motion a number of investigations are available [73, 8]. The same applies for the sensor density and topology [46, 73, 58].

### 4.5.3 *Results of the Parameter Investigation*

As described in Sec. 4.4.7 the results of the parameter investigation allow to derive a number of best practice considerations. But they also demonstrate the potential of tuning the Radio Sensor during runtime. Specifically, we showed that when iterating across carrier frequencies (frequency diversity) recognition performance improves. Hence, it seems worthwhile to consider additional parameters for continuous adaptation. The investigation showed that transmission power is the

most influential Radio Sensor parameter. By dynamically adapting transmission power, the radio signal illumination in the room and the influence of multipaths in the receiver are changed. Hence, many additional information about the monitored area could be acquired. As it is currently not considered as a parameter in device-free research, we suggest that future research investigates this parameter.

## 4.6 CONCLUSION

In this chapter, we identified three major aspects influencing radio signal measurements: subject, environment and Radio Sensor. For each aspect we described specific parameters. Five of these parameters were evaluated using the typical process chain of Activity Recognition using a total of 54 evaluation data sets. Finally, we ranked the parameters given their influence on the recognition accuracy. We found that the size of the environment is the most influential parameter, followed by the transmission power, the spacing of frequencies for frequency diversity, the number of frequencies and lastly the link length. In addition to identifying best practice considerations and novel insights for important parameters, the presented aspect-based perspective can foster a structured investigation methodology for building and validating suitable models for radio-based, device-free Activity Recognition in the Smart Home.

# 5 Holistic System Architecture for Practical Activity Recognition

This chapter in combination with Chapters 6–9 and Appendix A provides the validation of Thesis 3, demonstrating that a practical online Activity Recognition system can be realized. Specifically, this chapter is concerned with the analysis of requirements and the design of such a system. The chapter has the following structure: In the introduction we revisit the typical architecture of a device-free, radio-based sensor system. In the next section, we describe a software design pattern of an Inference System. Then follows an analysis of the requirements for practical Activity Recognition. Combining the findings of this chapter, a holistic system architecture is proposed.

## 5.1 Introduction

In Chapter 4 we have identified the two main components of a device-free, radio-based sensor system: The Radio Sensor and the Inference System (Fig. 17). The Radio Sensor measures radio signal characteristics which are utilized by the Inference System to derive context information.

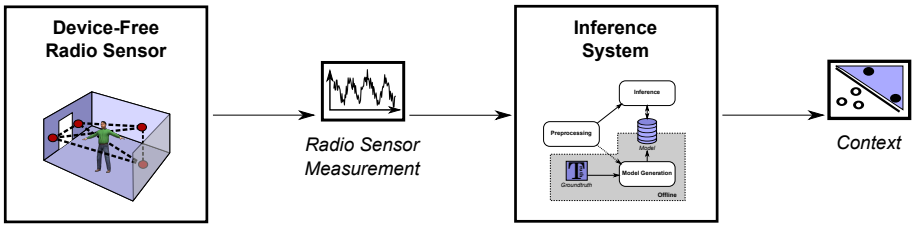



Figure 17.: Architecture of a device-free, radio-based Sensor System

The Radio Sensor and four of its parameters were investigated in Chapter 4. Appendix A describes its implementation. Hence, the focus of this chapter is the Inference System. Based on previous work, we derive a general software design pattern for radio-based Inference Systems independent of the context. We show that such a system can be reduced to three primary modules: *Preprocessing*, *Model Generation* and *Inference*.

Then we identify the three challenges which a practical system for online Activity Recognition using the device-free approach must address: Presence Detection,

Subject Discrimination and Activity Recognition. As explained in Section 5.3, the software design pattern of an Inference System is not suited to address all of these concurrently. Instead a holistic system architecture is proposed which integrates an Inference System for each specific challenge.

## 5.2 SOFTWARE DESIGN PATTERN FOR INFERENCE SYSTEMS

Previous device-free, radio-based recognition systems typically follow a common scheme [75]. Irrespective of the investigated context, calibration/training data are collected and used as a foundation for consecutive studies or the development of a novel system. Classic Activity Recognition follows a comparable approach [8]: First, sensor data containing the context is recorded, then an algorithm is trained and finally the algorithm is evaluated on previously unseen data. Thus, for an Inference System we propose an architecture based on the classical Activity Recognition approach adapted to the specifics of device-free, radio-based sensing (Fig. 18).

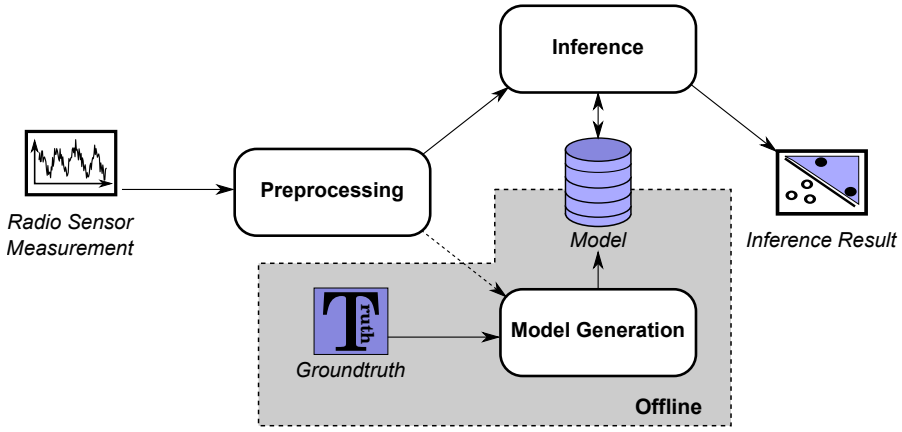



Figure 18.: General Inference System Design Pattern

An Inference System is composed of three submodules: *Preprocessing*, *Model Generation* and *Inference*. *Preprocessing* operates on *Radio Sensor Measurements* and applies transformations which emphasize sensor data characteristics related to the investigated context. It may filter the raw data or compute meta information (features) across a number of measurements. *Model Generation* is a submodule utilized in the offline phase of an Inference System. It uses *Groundtruth* data and possibly preprocessed measurement data to generate a *Model*. The *Model* is a description of how sensor data relates to the investigated context. *Groundtruth* provides an annotation of the preprocessed Radio Sensor data with the investigated context. The broken line between *Preprocessing* and *Model Generation* indicates that in some cases real world measurements from the Radio Sensor may not be necessary to build a *Model*. Finally, *Inference* uses the created *Model* to determine the context based on the current *Radio Sensor Measurements*. The *Inference* submodule may also change the *Model*, e.g. to tune it based on the *Inference Result*. The Inference System operates in two phases offline and online. In the offline phase the *Model*

is generated. In the online phase the *Model* is used for recognition and possibly further optimized.

## 5.3 REASONING FOR A HOLISTIC SYSTEM ARCHITECTURE

To validate Thesis 3, it is required to develop a practical online system. This means that the system can be deployed in a real world environment and that the produced outcomes are predictable. I.e. the evaluation of the system was performed such that the actual online performance can be estimated. This is crucial in order to show that even as the aspects (cf. Ch. 4) change over time recognition is still possible. Device-free Activity Recognition using signal strength measurements is a relatively new field. Previous research has largely conducted sample studies using cross-validation [73] or employed special radio signal magnitudes for analysis [82, 90].

In addition, there are two requirements which can not be addressed within a single such Inference System: presence and identity of a subject. Both are typically provided in classical wearable activity recognition and both are indispensable for practical device-free activity recognition. In the following two sections we consider and reason these requirements. To implement these requirements a complex holistic architecture is needed. This is described in the next section.

### 5.3.1 *Presence Detection*

To realize a robust radio-based Activity Recognition at first the detection of a subject is required. In Section 3 and Ref. [73] we have presented an initial investigation for device-free Activity Recognition. Therein eight nodes were installed in an office room of 4x5 m. The classes to be distinguished included six human physical activities and the class "Outside" which described the absence of a human in the monitored area. The cross-validation results of k-NN on the recorded testbed data were impressive (~90% accuracy). As pointed out in Chapter 3 one of the reasons that the performance was still lacking was the high confusion with the "Outside" class.

| Truth \ Predicted | Walking | Standing | Sitting | Sitting+Typing | Lying | Lying+Waving | Outside |
|---|---|---|---|---|---|---|---|
| Walking | 0.71 | 0.08 | 0.03 | 0 | 0 | 0.04 | 0.14 |
| Standing | 0.05 | 0.95 | 0 | 0 | 0 | 0 | 0 |
| Sitting | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Sitting+Typing | 0 | 0 | 0.04 | 0.96 | 0 | 0 | 0 |
| Lying | 0 | 0 | 0 | 0 | 0.92 | 0.07 | 0.01 |
| Lying+Waving | 0 | 0 | 0 | 0 | 0.16 | 0.83 | 0.01 |
| Outside | 0.01 | 0.02 | 0.02 | 0 | 0.05 | 0.02 | 0.88 |

Table 4.: Confusion Matrix for the k-NN on the Radio Sensor data

Inspecting the matrix replicated from Chapter 3 (Table 4), we find that the class "Outside" is frequently confused across all classes. Especially the low number of correct recognitions of the class "Walking" is associated with a large number of confusions with "Outside". The reasons which we find for the confusions with

"Outside" are discovered quickly: the test bed area is an office room which is only enclosed by gypsum walls. Radio propagation is only minimally affected by this walling. Hence, "Outside" activities such as subjects walking by influence radio links similarly as in-room activities. However, confusions are not only encountered for activities which include physical motion but also for activities with very limited motion.

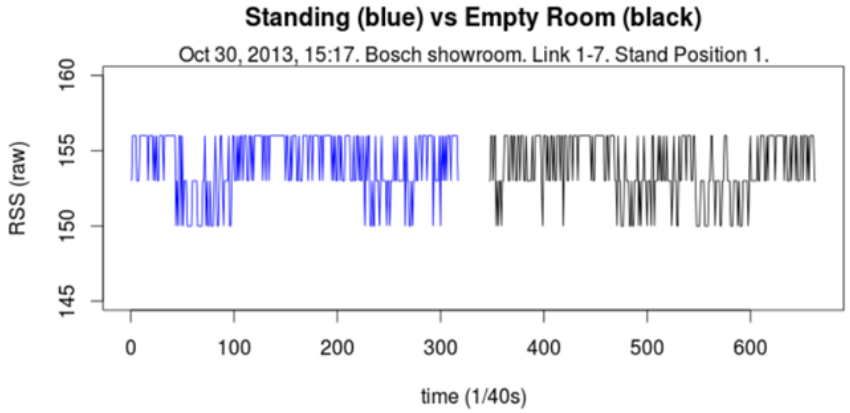



Figure 19.: "Standing" and "Empty Room" can produce similar RSS measurements

An illustrative example of this is shown in Figure 19 which is an extract from the data set described in Chapter 6. In this figure the received signal strength measurements for two activities on a single link are depicted. These are "Standing Still" and "Empty Room" (identical to "Outside" in the previous example). The figure shows that the raw data for both contexts are nearly identical with only little deviation in signal variance. However, also note that the signal only fluctuates between three distinct signal strength values which may also be induced by other environmental noise. Hence, especially in settings with limited spatial radio coverage of an area of interest it is likely that in-room activities induce similar patterns in RSS to out-room activities. This effect was also observed by Mrazovac et al. [55] in RSS and more recently by Wang et al.[90] in channel state information measurements.

Thus, to make device-free Activity Recognition practical it is mandatory to implement Presence Detection as an essential part of the overall system. But incorporating Presence Detection into the Inference System for Activity Recognition showed suboptimal results. Therefore this should be realized in an individual Inference System. We describe the implementation and evaluation of our Presence Detection System in Chapter 6.

### 5.3.2 *Subject Discrimination*

In Pervasive Computing, Activity Recognition has traditionally utilized contact-based sensing using accelerometers, gyroscopes and/or magnetometers. Shortcomings of this approach are battery-limited runtime and reduced comfort. However, it also comes with an advantage: the identity of the subject for which the

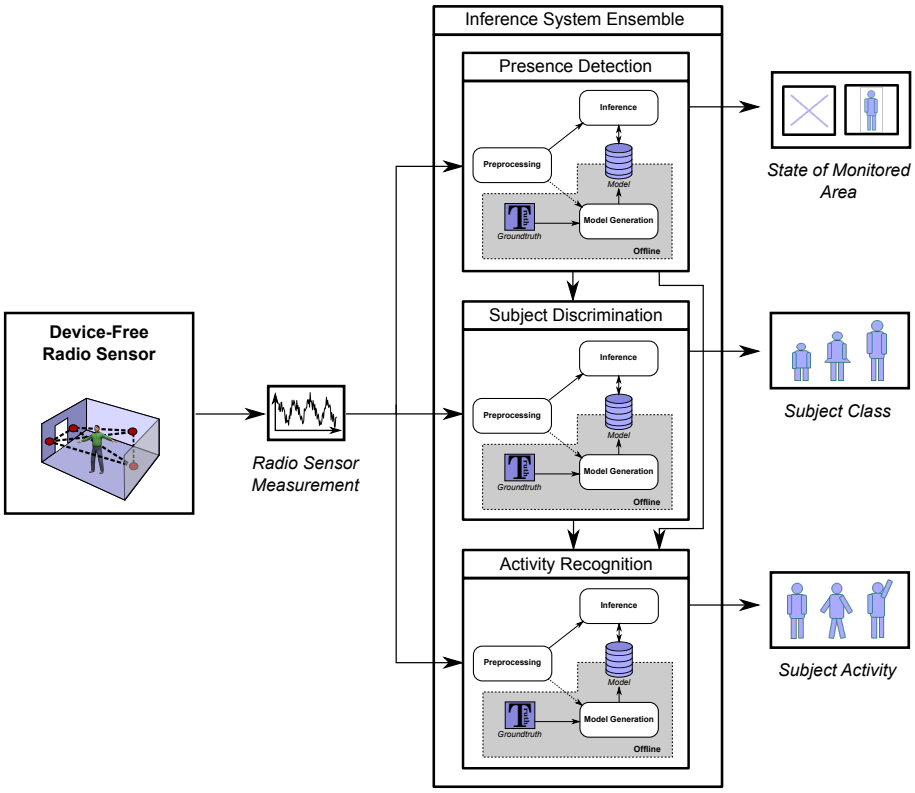



Figure 20.: Holistic System Architecture

activity is determined is known. This is an advantage for two reasons. First, it allows to relate recognized activities to a specific subject. This is important for implementing intelligent environments which adapt to the preferences of a specific user. Second, subject-specific recognition can improve accuracy. As subjects typically have different gaits or motion patterns this information can be used to improve true positive rates. Incorporating knowledge about a subject into an Activity Recognition system therefore increases usability and robustness. Hence, it is mandatory to consider Subject Discrimination as a crucial part of a device-free Activity Recognition system. It is likely that the features and preprocessing steps which are relevant for Subject Discrimination are not identical to those of Activity Recognition. Hence, it is advisable to implement Subject Discrimination in an individual Inference System. The implementation and evaluation of our Subject Discrimination system is detailed in Chapter 7.

## 5.4 HOLISTIC SYSTEM ARCHITECTURE FOR PRACTICAL ACTIVITY RECOGNITION

The previous section has motivated the need for Presence Detection and Subject Discrimination Inference Systems operating in cooperation with an Activity Recognition Inference System. The proposed holistic system architecture for a practical device-free, radio-based Activity Recognition system is shown in Figure 20.

The Radio Sensor continuously generates RSS measurements which are forwarded to each of the three Inference Systems within the ensemble. Each system has distinct specifically optimized *Preprocessing*, *Model Generation* and *Inference* modules and procedures. Each system further produces a classification result which may be used as is by an attached consumer e.g. a Smart Home appliance. Alternatively, the output of the systems can be combined to enrich the recognition results. This is indicated by arrows between the systems in the figure. Hence, the result of the Presence Detection is forwarded to the Subject Discrimination, while classification results from both systems are passed on to the Activity Recognition. In the following chapters we describe how each of the different subsystems was developed and evaluated.

## 5.5 CONCLUSION

In this chapter a general radio-based Inference System software design pattern was derived. Requirements for practical device-free, radio-based Activity Recognition were identified. We demonstrated that besides the actual recognition of activities also Presence Detection and Subject Discrimination must be addressed. Each of these contexts likely depends on other signal characteristics, hence incorporating them all into a single Inference System is unreasonable. Therefore, a holistic system architecture comprised of four subsystems was designed. In the general discussion of this dissertation, we integrate the systems into the proposed architecture and estimate the overall performance.

# 6 LoS-Cross: a Topology constrained Presence Detection System

This chapter supports the validation of Thesis 3: the development of a practical device-free Activity Recognition system. In Chapter 5, we reasoned the necessity of Presence Detection to implement practical radio-based Activity Recognition. Hence, the LoS-Cross Inference System presents the first building block in the proposed holistic system.

The chapter is structured as follows: In the introduction we reason the development of a new Presence Detection system and develop a working scheme for a new system. In the next section, we conduct a detailed analysis of existing systems. Then we describe the LoS-Cross architecture and experimentally evaluate a number of system configurations. The chapter finishes with a discussion and a conclusion.

## 6.1 INTRODUCTION

Current Presence Detection research can be categorized by the number of employed wireless transceivers. One approach is to employ a large quantity of transceivers (typically n>10) completely surrounding the area of interest [91, 40] use a large quantity of transceivers. Thus, the monitored area is well covered with wireless links ensuring that a subject will always cause a measurable change in RSS. The downside of such systems is the high installation effort. In contrast, other systems such as Ref. [45] or [103] work with only a few transceivers (typically n<5), possibly even on a pre-installed infrastructure. But these systems typically require that the subject is walking. A walking subject is equivalent to a stark environmental change which will eventually induce a measurable signal fluctuation. The downside of these systems is that in the case of a user activity with only little motion such as sleeping or sitting the system may fail to detect the presence. **Hence, current RSS-based systems can only detect presence if they employ a large number of transceivers or if the subject is continuously inducing a strong RSS fluctuation.**

In this chapter we present LoS-Cross, an alternative approach to Presence Detection. LoS-Cross is a bi-state system which relies on a defined transceiver topology

specifically optimized for Presence Detection. Within this topology it detects line of sight (LoS) crossings at the entrance to the area of interest to derive occupancy. As the room entrance is typically very narrow, the corresponding door link is short. Thus, a subject crossing the link will have a very strong impact blocking not only the line of sight path but interacting with a large number of paths between the transceivers. Hence, unsurprisingly LoS crossings can be detected with high reliability. However, it will be shown that similar signal fluctuations occur when subjects move inside or outside the room; potentially leading to a high number of false detections. As a countermeasure LoS-Cross employs a second link further away from the room entrance. Due to the distance between the link and the entrance it is only affected by motion inside the room. It will be shown that this approach not only helps to eliminate false detections but can also be used to determine the link crossing type.

To illustrate this work principle consider Figure 21. The figure shows the RSS for both links when a subject is entering (green triangles) or leaving (red squares) the room. The door link (top curve in the figure) shows multiple strong fluctuations similar to these two events. However, only "Entering" and "Leaving" are preceded or followed by a longer silence phase, i.e. a phase with very low variance, in the second link (lower curve in the figure). As the second link has typically very few changes during these periods we also refer to it as silence link.

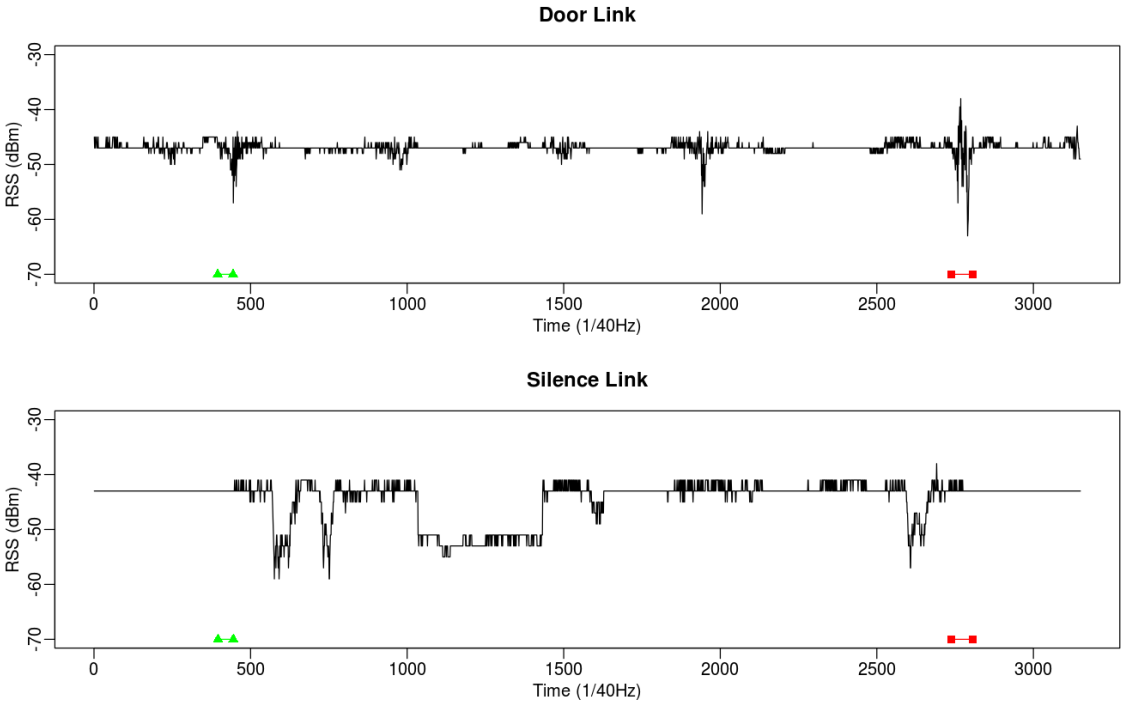



Figure 21.: "Entering" (between green triangles) and "Leaving" (between red squares) in the raw measurement data on both links

In this chapter we will evaluate LoS-Cross regarding different system parameters such as detection window size, suitable features and frequency diversity. In order

to quantify the impact of the silence link we will also evaluate the Presence Detection performance with and without this link. The system was evaluated using 14 evaluation datasets (and 14 corresponding calibration datasets) in a special testbed installation over a period of a week. We found that using the optimal configuration **LoS-Cross achieves a 92% accuracy which means that for 13 of 14 evaluation datasets it could correctly detect the exact time of presence of a subject independent of the subject's activity**.

## 6.2 RELATED WORK

The first publications relevant to Presence Detection appeared as early as 2007. Since then a growing corpus of research has considered this challenge. We will review these in the following subsections.

### 6.2.1 *Presence Detection using few Nodes*

The first device-free, radio-based motion detection system was published by Youssef et al. [99] in 2007. The Radio Sensor consisted of two access points (APs) and two wireless clients (MPs) installed in the corners of a rectangular room. The APs sent WiFi beacons with 10Hz. Each beacon was received by the two MPs yielding a total of four wireless links across the room. Experimental evaluation had a subject walk in the room, stop at each of four locations for 60 s, and then leave the room. The moving average and moving variance were tested as detection features. The variance was found to be superior. 100% accuracy was achieved using a 4 s variance window and at least three links.

May et al. [53] continued research into device-free Presence Detection by testing the motion detection system of their group [99] in a more realistic environment (a lively office). The former system was tested in this uncontrolled environment (square office room of 5x5 m) and performance decreased to 80%. They succeeded to improve the system performance by 10% using a maximum likelihood estimator but make no statement about the actual experiment/data collection.

Yang et al. [97] presented an intrusion detection system which they evaluated using different link topologies: a single link, two links, and a three links each in a rectangular tunnel-like topology. They conducted a diverse set of experiments involving standing and walking activities of an intruder. However, all activities are conducted such that they either cross the LoS (walking) of a link or are conducted directly in LoS ensuring a strong effect on the signal, making them easier to detect. Using empirical thresholds on 10 s-long sliding windows with mean and variance as features they achieve accuracies of 90% and 100% using two or three links, respectively.

RASID [45] is a WiFi-based Presence Detection and localization system which uses statistical anomaly detection while adapting to changes in the environment to provide accurate long-term detection. The system has a single calibration phase in which it determines silence thresholds i.e. normal signal parameters when the

environment is empty. In a monitoring phase the system collects signal strength readings to decide if human activity is present. Detection is realized by checking each link for anomalies. An anomaly occurred when a sliding variance for a link exceeds the calibration threshold. A special feature is a refinement module in which the system adapts the calibration thresholds in order to keep detection stable even over longer time periods. The system was evaluated in two testbeds using a number of evaluation datasets. A single evaluation set comprised a single subject walking continuously through the monitored area (floor of a building). The achieved detection accuracy was 93% (F-measure) for both test beds if the system was evaluated directly following calibration. And 92% (F-measure) when the system was trained two weeks before the evaluation. Test bed one was a complete office floor of 185 $m^2$ with four access points and three WiFi clients distributed equally over the office floor. Test bed two was a two-floor home building of 140 $m^2$. RASID has the advantage that it could be applied directly to todays offices and homes. However, it will only detect presence of a moving subject.

Mrazovac et al.[55] analyzed IEEE 802.15.4 received signal strength (RSS) using thresholds on the eigenvalues of the principal components of RSS of multiple links to derive human presence. Evaluation took place in an office room with gypsum walls and an area of 5.36x5.30 m using four wireless transceivers. The topology was irregular i.e. not all nodes were located in a corner of the room. For training and test only a single relatively short data set was analyzed and evaluated using principal component analysis (PCA). The performance on this test set was 98% with PCA over 25 s windows. However, it is difficult to estimate the actual performance as no independent training and test sets were used in the evaluation.

### 6.2.2 *Presence Detection using many Nodes*

Kaltiokallio and Bocca [40] presented a system for intrusion detection and tracking. Presence of an intruder was detected in a rectangular area surrounded uniformly by a large number of IEEE 802.15.4 nodes (n=16). In their system, Presence Detection is based on an empirically estimated threshold and the comparison of short and long sliding variance windows. The variance is calculated over RSS values filtered using a weighted moving average filter. The algorithm was tested in four test beds covering 16 $m^2$, 36 $m^2$, 36 $m^2$ and 64 $m^2$. All test beds were installed in a larger area or foyer without walls close to the transceivers. Achieved accuracies for Presence Detection ranged from 19% (small area) to 14% (large area).

Wilson and Patwari [91] patented a system utilizing many nodes surrounding the area of interest e.g. a house in a regular topology. They detect presence and motion based on calibration using metrics of dispersion such as variance for each of the links and provide motion detection by comparing these calibration data to actual measurements during an online phase. In their description they also include an online calibration example using mean differences similar to Youssef et al. [99].

### 6.2.3 *Presence Detection using specialized Radio Hardware*

Zhou et al. [103] presented a human detection system using channel state information (CSI) from adapted WiFi drivers with a single link. CSI provides a finer-grained temporal and spectral information of a wireless 802.11n link but is currently not commonly available. The authors analysed detection performance using a fingerprinting approach for a subject standing and walking in close proximity (0.5 m, 1.0 m) in a circle around the receiver. The experiment was repeated in two environments (multipath rich and multipath scarce). The system achieved an average detection rate of 90%. Besides the use of CSI, this approach would need a very high distribution of wireless receivers for room based Presence Detection (1.0 m spacing).

Wang et al. [90] presented the E-Eyes system an activity "identification" system utilizing CSI. The system employs a recognition scheme in which activities which can not be identified directly are estimated through a priori detection of the currently occupied room, e.g. cooking will take place only in the kitchen. Interestingly, to detect the occupied room E-Eyes utilizes walking identification followed by doorway passing detection. The average detection accuracy of doorway passing and identification using CSI was 96.25% with two monitored doorways in each of the two test beds. The system utilized three links, the sliding variance and the earth movers distance to identify the activity and the crossed doorways.

## 6.3 THE LOS-CROSS SYSTEM

LoS-Cross is a calibration-based device-free, radio-based Presence Detection system. Its approach differs from previously reviewed systems in two main aspects: Firstly, the detection is based on strong RSS fluctuations in the link spanning the room entrance which are filtered using a second link in combination with a room occupancy state-model. Secondly, the approach requires a specific two link Radio Sensor topology which supports this detection approach. In this section, we present the system architecture and detail the system modules and their implementation options. These options are evaluated in Section 6.4.

### 6.3.1 *Architecture Overview*

The architecture of LoS-Cross is shown in Figure 22. It builds on the general Inference System Design pattern (cf. Ch. 5). LoS-Cross operates in a calibration (offline) and recognition (online) phase.

For calibration a subject enters the room and walks until he reaches the far corner of the room (from the door). During this time RSS measurements from the Radio Sensor are forwarded to the *Feature Calculuation* module. In this module a feature is computed on a window of RSS values for each link and each frequency. The computed feature vector is passed to the *Calibration* module. *Calibration* has two submodules which operate on the complete calibration dataset. The *Door Link*

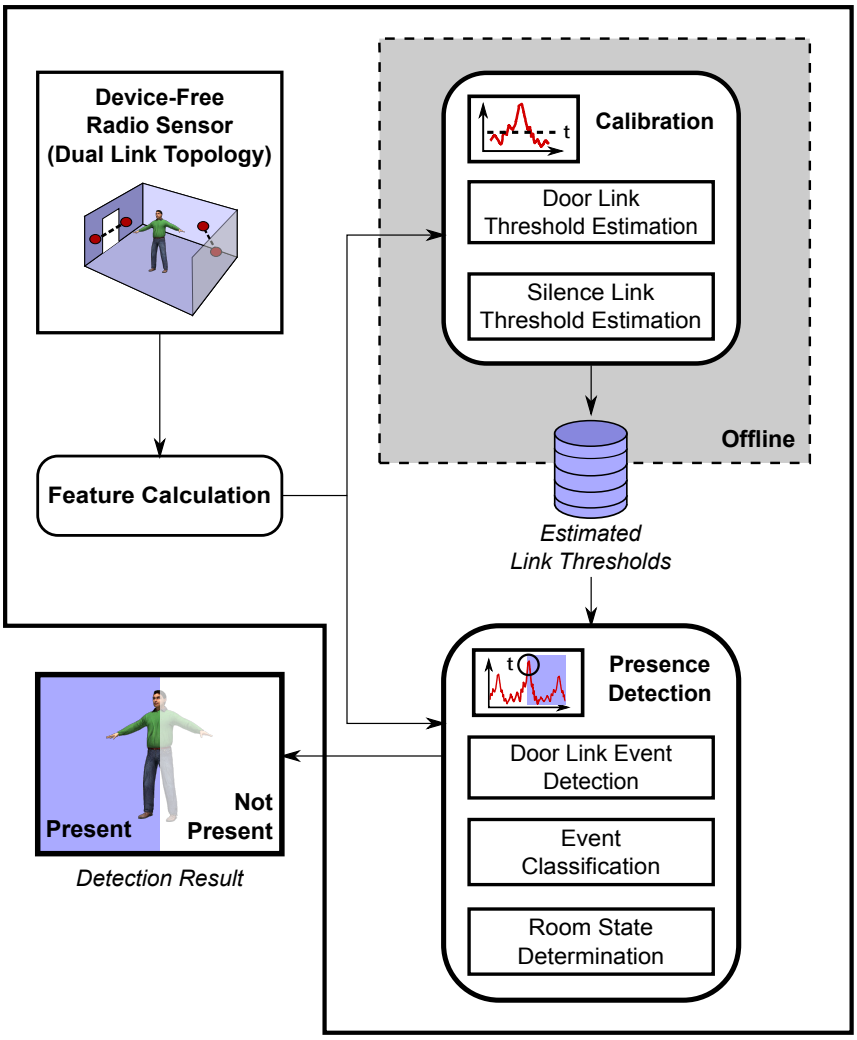



Figure 22.: LoS-Cross System Architecture

*Threshold Estimation* estimates the minimal threshold indicating a subject crossing the entrance link. The *Silence Link Threshold Estimation* estimates the maximal fluctuation of the link in the far end of the room *before* the subject entered the room, i.e. before the strongest observed fluctuation in the door link.

In the online phase the data from the *Radio Sensor* is continuously streamed through the *Feature Calculation* into the *Presence Detection* module. In its *Door Link Event Detection* submodule the door link data are continuously analyzed. If a crossing was detected the *Event Classification* analyzes the preceding (subject entering) or subsequent data (subject leaving) of the silence link. If these are below the silence threshold the *Room State Determination* is triggered. It outputs the LoS-Cross detection result, i.e. whether or not a subject is present in the room.

### 6.3.2 *Feature Calculation*

LoS-Cross is built upon the observation that a subject moving through the line of sight of a link induces very strong fluctuations (cf. Fig. 21). The task of *Feature*

*Calculation* is emphasizing this fluctuation to simplify algorithmic detection. In statistics, measures of dispersion such as variance and standard deviation are employed to indicate strength of fluctuation within a sample. Previous Presence Detection research [99, 45] also successfully employed these measures. Hence, we investigated the sample standard deviation std(X).

$$\mathrm{std}(X) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - x_\mu)^2} \tag{10}$$

Where $N$ is the number of elements of $X$ and $x_\mu$ is the mean of $X$.

Further on, we investigated the range ($\mathrm{rng}(X)$). The range is more strongly affected by outliers (as which we may interpret human induced signal fluctuations) in $X$ than the standard deviation:

$$\mathrm{rng}(X) = \max(X) - \min(X) \tag{11}$$

Another aspect of the feature calculation is the size of the window used for the feature calculation. In general, sliding windows seem advisable as signal fluctuations may be missed if they take place between two windows. Previous Presence Detection research also achieved good results using sliding standard deviation windows. In Ref. [99], Youssef et al. reported that a sliding standard deviation window of 4 s (sample rate: 10 Hz) delivered the best results. But in contrast to their work, LoS-Cross focusses on detecting link crossing only. If a window is 4 s it is likely that it may also include motion patterns of the period before or after entering the room, blurring the actual influence of the crossing. Additionally, if we consider a brief calibration process ("Walking into the room") a 4 s window might simply be too long.

Clearly, the optimal window size is a function of the subject's speed and size and the expansion of the most influential area of the link. The most influential area can be described by the first Fresnel zone which includes the line of sight (LoS) signals between transmitter and receiver [63]. If the LoS path is obstructed the effect on RSS is significant because it is the shortest path and therefore contributes the largest amount of energy to the received signal. Thus, obstructions lead to high signal fluctuations at the receiver. The first order Fresnel zone is an ellipsoid with foci in the transmitter and receiver. We can use the Fresnel equation (Eq. 12) to compute the maximal expansion of this ellipsoid for a door link of 1.20 m.

$$F_n = \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}} \tag{12}$$

In this equation, $n$ relates to the order of the Fresnel zone. $\lambda$ is the wavelength and $d_1$ and $d_2$ are the distances from receiver and transmitter to the point for which the radius of the Fresnel ellipsoid is computed. Using $n = 1$, $\lambda = 0.01228$ m (2440 MHz) and $d_1 = 0.56$ m, $d_2 = 0.56$ m we find that the diameter of the doorway

Fresnel zone is 0.38 m. Considering a preferred walking velocity of 1.42 m/s [14] and an average sagittal abdominal diameter (waist diameter) of 0.22 m [39], the average time for passing through the Fresnel zone can be estimated as 0.422 s. Thus, we will evaluate sliding window sizes $w_s$ around this average passing time as follows: $w_s = 0.125\,\text{s}, 0.250\,\text{s}, 0.425\,\text{s}, 0.500\,\text{s}, 1.000\,\text{s}, 2.000\,\text{s}$. Given a time series of RSS recordings ($X$) we compute the new time series of features ($Y$) by applying the feature function f (range, standard devation):

$$Y_t = \text{f}(X_t, \ldots, X_{t+w_s}) \tag{13}$$

### 6.3.3 *Calibration*

The *Calibration* module estimates the link thresholds using a calibration data set. The *Door Link Threshold Estimation* submodule estimates a single or multiple thresholds for crossing the room entrance depending on the number of employed carrier frequencies in the Radio Sensor. The *Silence Link Threshold Estimation* uses these thresholds to derive silence thresholds which are defined by the signal levels occurring before the subject enters the room.

#### *Door Link Threshold Estimation*

Based on the *Radio Sensor* configuration a link may be measured on a single or multiple frequencies. We consider the single and three-frequency case here.

SINGLE FREQUENCY OPERATION For a single frequency system the threshold is estimated by finding the maximal fluctuation in the feature time series $Y_t$ of the calibration data set. As the maximal value may be encountered multiple times, the first and last $(t_{\text{start\_single}}, t_{\text{end\_single}})$ occurrence of the maximum in the time series is extracted. To make the detection more sensitive this window is expanded to the left $(t_{\text{start\_single}} - 1)$ and right side $(t_{\text{end\_single}} + 1)$. The minimum over the whole range then constitutes the detection threshold ($\text{thr}_{\text{cross\_single}}$). Figure 23 illustrates the process. In the figure the gray box indicates the window covering $t_{\text{start\_single}} - 1$ to $t_{\text{end\_single}} + 1$ over which the threshold (broken line) is defined. Along with the detection threshold, the size of the threshold window (r1) and the time span (r2) from the start of the recording until the beginning of threshold window is stored.

$$\text{thr}_{\text{cross\_single}} = \min(Y_{t_{\text{start\_single}}-1}, \ldots, Y_{t_{\text{end\_single}}+1}) \tag{14}$$

$$\text{r1}_{\text{single}} = t_{\text{end\_single}} - t_{\text{start\_single}} \tag{15}$$

$$\text{r2}_{\text{single}} = t_{\text{start\_single}} - 2 \tag{16}$$

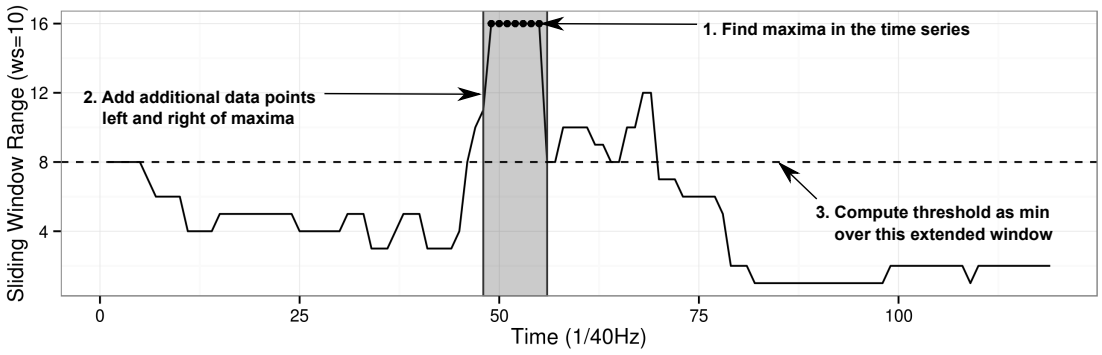



Figure 23.: Illustration of threshold estimation on feature data of a single frequency calibration data set.

TRIPLE FREQUENCY OPERATION For a system operating on three frequencies the maximal fluctuation for each of the individual frequencies is computed. In the detection module we assume that all frequencies react at almost the same time when a subject obstructs the LoS. Thus, in the next step we determine the first ($t_{\text{start\_triple}}$) and last occurrence ($t_{\text{end\_triple}}$) of the maxima over all frequencies of the door link. Finally, we determine the minimum value in this window for each frequency as the frequency specific detection thresholds (cf. Eq. 17, 18, 19). Figure 24 illustrates the process of triple frequency threshold estimation. As in the previous figure the highlighted points indicate the maxima per frequency. The gray box shows the global range of the maxima i.e. from $t_{\text{start\_triple}}$ to $t_{\text{end\_triple}}$. The broken line shows the frequency threshold based on the minimum encountered in this range. As in the single frequency case we also store window size (r1) and the number of samples from the start of the calibration recording until the first maxima (r2). Together with the three frequency thresholds the values computed are as follows:

$$\text{thr}_{\text{cross\_triple\_f1}} = \min(\text{F1}_{t_{\text{start\_triple}}}, \ldots, \text{F1}_{t_{\text{end\_triple}}}) \tag{17}$$

$$\text{thr}_{\text{cross\_triple\_f2}} = \min(\text{F2}_{t_{\text{start\_triple}}}, \ldots, \text{F2}_{t_{\text{end\_triple}}}) \tag{18}$$

$$\text{thr}_{\text{cross\_triple\_f3}} = \min(\text{F3}_{t_{\text{start\_triple}}}, \ldots, \text{F3}_{t_{\text{end\_triple}}}) \tag{19}$$

$$\text{r1}_{\text{triple}} = t_{\text{end\_triple}} - t_{\text{start\_triple}} \tag{20}$$

$$\text{r2}_{\text{triple}} = t_{\text{start\_triple}} - 1 \tag{21}$$

In Eq. 17, 18 and 19 the feature time series data for the three frequencies is denoted by F1, F2 and F3, respectively.

*Silence Link Threshold Estimation*

The purpose of the silence link is to reduce false positive crossing detections and to determine the type of the crossing event i.e. "Entering" or "Leaving". The intuition behind the silence link is that a link most distant from the entrance should be

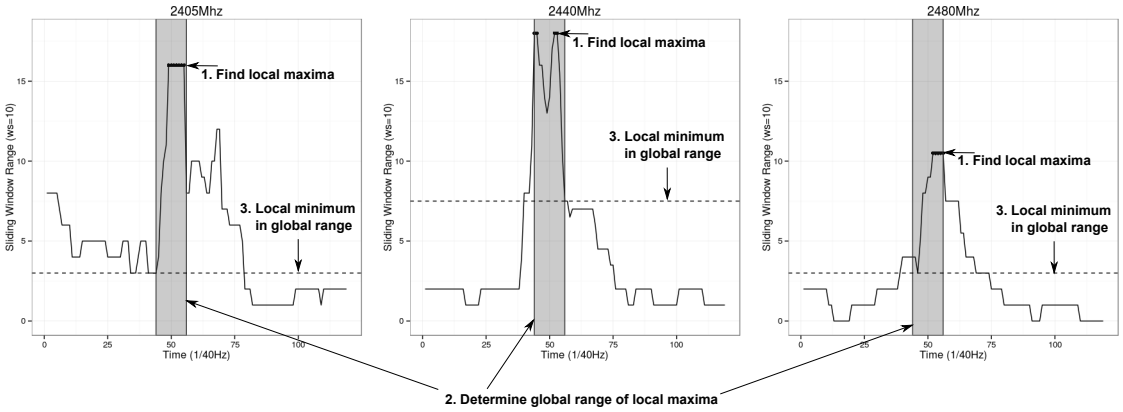



Figure 24.: Threshold estimation in triple frequency operation

nearly unaffected by motion *in front of* the door to the room prior to entering. If this assumption holds, we can classify an event as "Entering" if the silence link is static before the event. Vice versa if the silence link is static after a crossing event it may be classified as "Leaving". If non of the criteria is fulfilled an event is eliminated as a false positive detection. Thus, in contrast to the door link we need to identify a silence threshold for this link. Before determining the threshold(s) the time series data of the silence link is passed through the same feature computation as the door link data. In the following paragraph we explain the determination of silence state thresholds for single and triple frequency operation.

SINGLE FREQUENCY OPERATION The silence threshold ($\text{thr}_{\text{silence\_single}}$) is the maximal value on the feature time series of the silence link ($Y$) in the window between start of the calibration recording and before the time stamp of the first element of the door threshold window ($t_{\text{start}} - 1$).

TRIPLE FREQUENCY OPERATION When operating on three frequencies the thresholds ($\text{thr}_{\text{silence\_triple\_f1}}$, $\text{thr}_{\text{silence\_triple\_f2}}$, $\text{thr}_{\text{silence\_triple\_f3}}$) are defined in the same way as in the single frequency but for each frequency individually. Thus, for each frequency the maximum encountered in the corresponding time series F1, F2 or F3 before $t_{\text{start\_triple}}$ is set as threshold.

### 6.3.4 *Presence Detection*

This module is operated in the online phase to perform the subject Presence Detection. The module operates on continuous input from the *Feature Calculation* to detect a possible crossing event via the *Door Link Event Detection* which is then either discarded or identified as "Entering" or "Leaving" using the *Event Classification*. If the event has not been discarded as false positive it is passed over to the *Room State Determination* module. Depending on the newly detected event and the internal state of the *Room State Determination* the room state is changed.

*Door Link Event Detection*

Link crossing detection is performed by comparing the *Estimated Link Thresholds* to the currently incoming feature data for each link and frequency. Depending on the used frequency mode the detection algorithm is performed differently. In single frequency operation a crossing is detected if the feature data coming from the *Feature Calculation* is greater or equal than the calibration threshold. In triple frequency operation a crossing event is detected if the incoming feature data on **at least two frequencies** for the door link is greater or equal to the calibration thresholds for these frequencies.

*Event Classification*

The detected events from the *Door Link Event Detection* are the input to this module. The silence detection investigates the fluctuation of the silence link in relation to a detected crossing event. Thus, it evaluates whether the silence link stays below the calibration threshold for some time period before or after the event. If this is not the case the event is discarded as false positive. Note that we use the time spans (r1, r2) discovered in the *Calibration* module as silence time periods. If the silence threshold is met prior to the event, the event is classified as "Entering". If the silence threshold is met following the event then it is categorized as "Leaving". In single frequency operation the silence threshold is only compared to the calibration threshold on a single frequency. In multi frequency operation the type of an event is decided based on majority voting against the silence threshold over all three frequencies.

As an example we give the formal description of the crossing event validation and categorization for the single frequency case:

$$\text{test}(t_{\text{event}}) = \begin{cases} \text{Entering} & \text{if } Y_{t_{\text{event}}-(\text{r1}_{\text{single}}+\text{r2}_{\text{single}}),\ldots,t_{\text{event}}-\text{r1}_{\text{single}}} \leq \text{thr}_{\text{silence\_single}} \\ \text{Leaving} & \text{if } Y_{t_{\text{event}}+\text{r1}_{\text{single}},\ldots,t_{\text{event}}+\text{r1}_{\text{single}}+\text{r2}_{\text{single}}} \leq \text{thr}_{\text{silence\_single}} \\ \text{False Positive} & \text{else} \end{cases}$$

Note that the condition for leaving introduces a delay of $\text{r1}_{\text{single}} + \text{r2}_{\text{single}}$ into the system. In the special case an event is detected as "Entering" and "Leaving" at the same time the type is decided based on the internal state of the *Room State Determination*:

$$\text{test}(t_{\text{event}}) = \begin{cases} \text{Entering} & \text{if test}(t_{\text{event}}) == \text{Entering \&} \\ & \quad \text{test}(t_{\text{event}}) == \text{Leaving \&} \\ & \quad \textit{Room State Determination} == \text{Not Present} \\ \text{Leaving} & \text{if test}(t_{\text{event}}) == \text{Entering \&} \\ & \quad \text{test}(t_{\text{event}}) == \text{Leaving \&} \\ & \quad \textit{Room State Determination} == \text{Present} \end{cases}$$

*Room State Determination*

In this module the actual room occupation state is derived. Internally, this module employs a state machine with the two states "Present" and "Not Present". When the system is switched on it will start in the state "Not Present" and wait for a detection event to change the state. If the systems receives an "Entering" event it will change to the state "Present". However, if it receives "Leaving" events in the state "Not Present" the state will not be changed. Likewise if the system is in the state "Present" it will not change state until it has received a "Leaving" event. Figure 25 illustrates the behavior of the state machine.

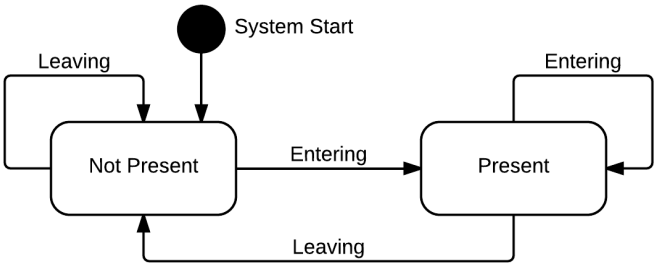



Figure 25.: State machine in the *Room State Determination* module

In the evaluation section we will also investigate how the system behaves when the silence link is not used i.e. if the *Event Classification* is not available. In this case no information regarding the type of event is available and the internal state machine switches between states as soon as an event is detected.

### 6.3.5 *Device-Free Radio Sensor in Dual-Link Topology*

The Radio Sensor is an arrangement of wireless transceivers installed in the room of interest. It acquires RSS between all transceivers and relays them to the *Feature Calculation*. LoS-Cross constrains the Radio Sensor installation to a specific topology. In particular, LoS-Cross requires two links installed in specific locations of the room. One link must cross the entrance and the second link must span a part of the room far away from the entrance. The knowledge which data belongs to which link needs to be configured in the LoS-Cross system.

## 6.4 EXPERIMENTAL EVALUATION

In this section we present the evaluation of the LoS-Cross system. First, we describe the used testbed and experimental execution. Next, we evaluate the system performance in respect to different system parameters. The section concludes with a summary of the best configuration and a comparison of our results to existing research.

### 6.4.1 *Testbed and Data Collection*

The system is installed in an office room of 5.50x4.60 m and a height of 3.05 m. Three walls of the room were constructed of gypsum boards, while the outside wall was made of concrete. The room had two windowed walls and a concrete ceiling. The room was outfitted with typical office pressboard furniture and multiple workstations. The width of the doorway is 1.10 m.

The Radio Sensor was installed using four Jennic 5139 IEEE 802.15.4 2.4 GHz transceiver nodes. Two nodes were installed left and right of the entrance yielding a door link length of 1.20 m. The stabilizing silence link was located in the far left corner from the room entrance. This link was 3.27 m long. All transceivers were installed in 0.80m height. Transceivers measured RSS on both links at 40 Hz on three carrier frequencies ($f_c = 2405$ MHz, 2440 MHz, 2480 MHz).

The system was evaluated over the time frame of one week. All experiments were conducted by the same subject (male, 1.76 m and 72 kg). During this week calibration and evaluation data sets were recorded multiple times on four days. For each evaluation data set a calibration data set was recorded beforehand. A total of 14 recordings (each containing a calibration and an evaluation data set) were created. Figure 26 shows the installation of the transceivers as well as the experimental setup.

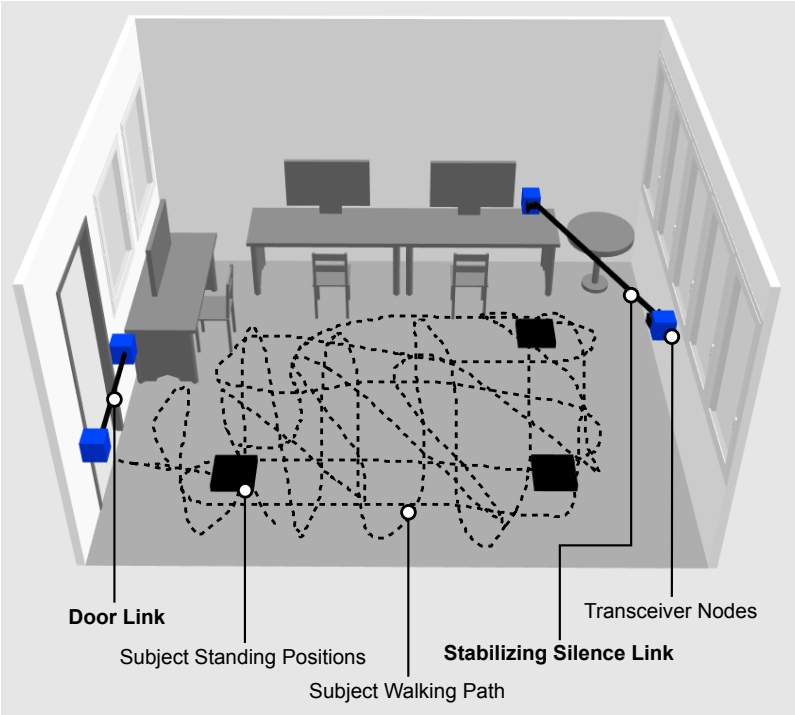
#### *Calibration*

For calibration the subject walked once from outside to the far left corner of the room. The recording was started when the subject was one meter in front of the door. During calibration raw RSS values for the two links were acquired by the *Radio Sensor*, passed on to the *Feature Calculation* and analyzed by the *Calibration* module to determine the activity and silence thresholds and ranges for the two links. In order to compare single frequency and triple frequency operation, the system computed both, the single frequency thresholds (for each of the three frequencies) and the combined thresholds and ranges as previously described. The average duration of a calibration data set is 3.2 s.

#### *Evaluation*

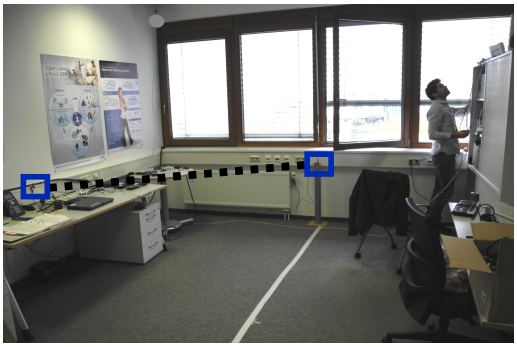
The evaluation data set was recorded directly following the calibration set. Evaluation involved the following sequence:

1. Standing still in front of the room for 5 s
2. Entering the room and walking to a standing position in the far left of the room
3. Standing at this posiiton for 10 s
4. Walking in the room for a while and stopping at the position at lower right of the room
5. Standing at this position for 10 s




(a) Schematic of the Testbed and Experiment

(b) Image of Door Link Transceivers



(c) Image of Silence Link Transceivers

Figure 26.: Testbed for Experimental Evaluation

6. Walking in the room for a while and stopping at the position close to the door
7. Standing at this position for 10 s
8. Leaving the room
9. Stands still in front of the room for 5 s

The different standing positions are highlighted as black squares in the Figure 26a. The average duration of a single evaluation data set is 2.5 min.

### 6.4.2 *Evaluation Metric*

LoS-Cross is a presence or occupancy detection system. Thus, recognition is correct only if subject presence is detected during the **complete time** the subject is

in the room. Classification is incorrect if the presence phase of the subject was not predicted **exactly**. Thus, if the subject is detected to leave the room earlier/later or enter the room earlier/later than recorded in the groundtruth the detection accuracy is 0%:

$$\text{Accuracy(testset)} = \begin{cases} 100\% & \text{period of subject presence exactly detected} \\ 0\% & \text{else} \end{cases}$$

### 6.4.3 *System Performance for Different Parameter Settings*

In this section we present the system evaluation. We investigate the effect of the sliding window size, feature, single frequency vs triple frequency operation and the performance gained when utilizing the silence link. We will investigate one parameter at the time keeping default values for all other parameters. The default values are shown in Table 5.

| **Parameter** | **Default Value** |
|---|---|
| Radio Sensor Operation | Triple Frequency |
| Feature Function (f) | Range (rng) |
| Feature Window Size | 0.25 s |
| Used Links | Door Link+Silence Link |

Table 5.: Default values for the different LoS-Cross parameters.

*Influence of Window Size*

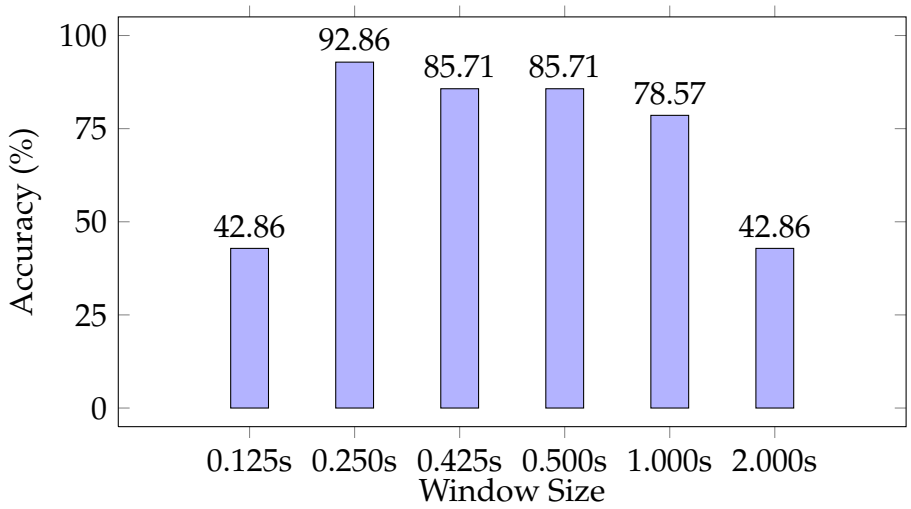



Figure 27.: Achieved accuracies using different window sizes

Figure 27 shows the achieved accuracies using the different window sizes defined in Section 6.3.2. The figure shows that a window size of 0.25 ms outperformed

all other configurations, detecting the subject's exact time of presence in 13 of 14 recordings. Window sizes of 0.42 s and 0.50 s still detected the phase of presence of a subject in 12 of 14 cases correctly. Shorter or longer windows lead to a decreased detection performance.

The decreased performance in longer windows seems to contrast with observations by Youssef et al. [99] who report much longer windows (4 s) to provide better results. However, recall that we are not seeking to detect motion to infer subject presence but the single crossing of the LoS which is typically a short event with strong fluctuation. Hence, for this detection task an increased window length rather increases the chance of false detection. For instance, if a subject incurs a high RSS range simply by moving to a different location within the room.

The decreased performance for the very short windows of 0.125 s can be explained with the limited number of samples (5) which are considered in this case: As the subject has a limited velocity the duration of the induced signal fluctuation is longer than 0.125 s. Hence, the complete range inflicted by an "Entering" or "Leaving" event is not captured in only 5 samples.

With this result we find our assumption confirmed, showing that for link crossing detection relatively short window sizes provide better results. Nevertheless, the anticipated best window length of 0.425 s only provided the second best result, indeed the optimal window length is only half the expected length. Following our deduction in Section 6.3.2 we can give three possible reasons for this result:

1. The subject moved with approximately twice the preferred walking speed.
2. The subject's sagittal waist diameter was considerable smaller than the average.
3. The area of influence of the link was smaller than estimated or considerably changed by obstructions in the near field of the transceivers (walls, door frame, door).

Walking speed was not measured in our experiments. However, walking speed would need to be doubled to explain the discovered window size. As the experimenter was instructed to walk with normal pace this is unlikely as sole explanation. The waist diameter of the subject measured a posteriori was 0.20 m which also does not explain the strong deviation. In fact, employing the Fresnel zone model the waist diameter would have to be smaller than zero to explain the discovered window size. Therefore we believe that the most likely reason is a distorted link area of influence. Fresnel zone computation assumes free standing transmission towers while our transceivers were very close to the walls and partially covered by the opened door.

*Influence of Feature*

Previous device-free presence research has shown the superiority of dispersion features over features of central tendency due to their temporal stability and sensitivity [45]. While the mean is affected by environmental changes, dispersion features, which quantify the spread of the distribution, are less prone to such

effects. Building on these findings we selected range and standard deviation for the feature investigation.

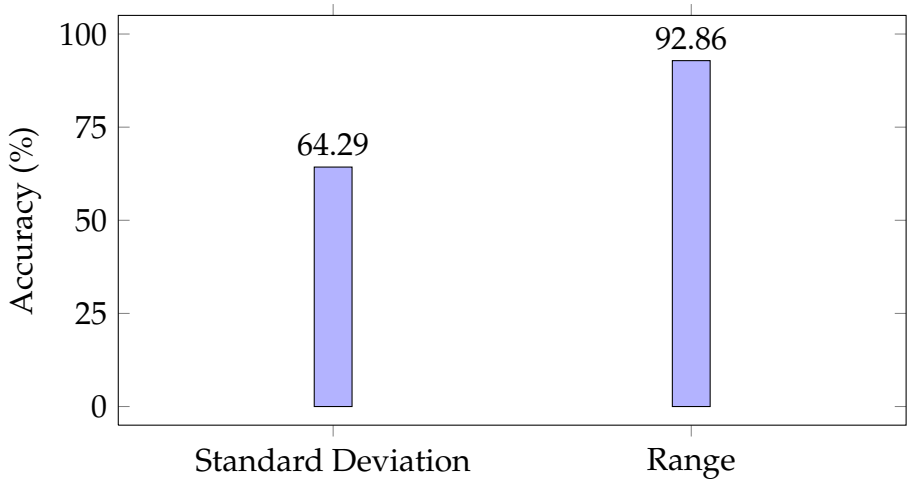



Figure 28.: Achieved accuracies using different features

Figure 28 shows the performance of the standard deviation feature vs the range feature. The range outperformed standard deviation by 30% points. We ascribe this result to the fact that single short fluctuations typically indicate a crossing event. This is well captured in the range feature. However, in very noisy environments this could lead to many abnormal measurements and therewith to a higher chance of false detections.

On the other hand, in the standard deviation a single peak will be related to the distribution of all values in the feature window, reducing its impact. Also consider that the standard deviation of a distribution with a number of medium fluctuations can be similar to or larger than a distribution with only a single strong fluctuation. This also increases the likelihood of confusing in-room induced fluctuations vs crossing fluctuations. In contrast, the range feature will yield distinct values for these two windows.

Hence, while the range feature showed superior performance in our experiments one future research direction might be the investigation of filtering approaches and feature combinations for improved detection performance and robustness in more noisy environments.

*Influence of Frequency Plurality*

Figure 29 shows the accuracy of the mean over the single frequency performance vs the performance of fusing the information from all three frequencies using majority voting. Clearly, the multi frequency operation outperformed any of the single frequencies. This is not surprising as the subject has a different influence on each of the frequencies. Combining the information increases the chance to detect crossing events. However, the triple frequency operation also uses a larger threshold estimation window. While this initially decreases the thresholds leading to a more sensitive crossing detection, the majority voting in both the *Door Link*

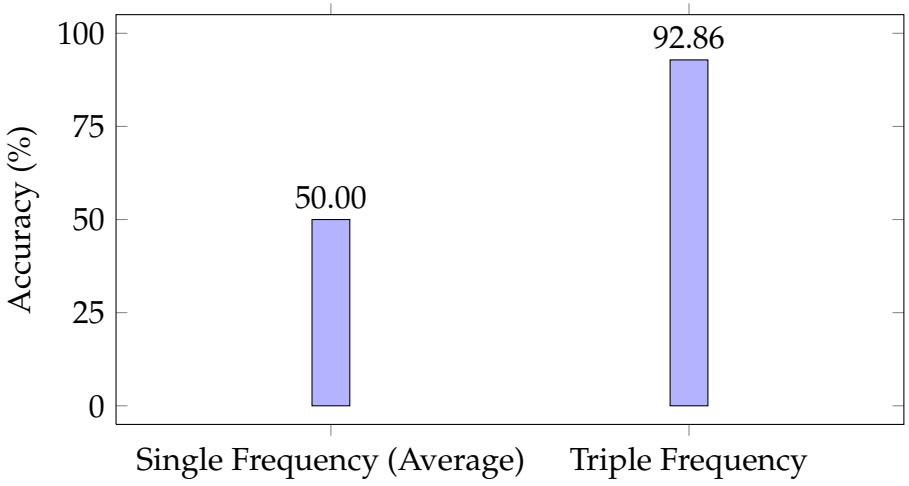



Figure 29.: Achieved accuracies when using a single frequency or three frequencies

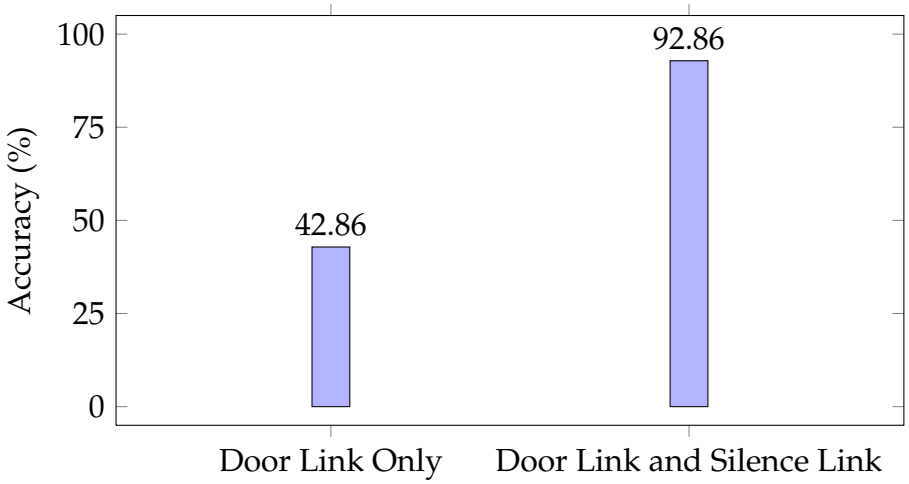



Figure 30.: Achieved accuracies when using the door link only and when using the combination of door and silence link

*Event Detection* and *Event Classification* stabilizes the detection and leads to an improved performance over the single frequency case.

*Influence of Silence Link*

Figure 30 shows the accuracy with/without the silence link. When the silence link information is omitted the unfiltered door crossing events are relayed to the *Room State Determination*. Looking at the figure the importance of the silence link becomes obvious. Without this link performance degraded strongly to only about half. This means that a lot of in-room events e.g. walking in close proximity to the door induced link fluctuations exceeding the calibrated thresholds. The silence link has a very strong filtering function removing these false positives.

Figure 31 illustrates the observation more closely on a complete evaluation set after system calibration. The figure shows the time series data of all three

frequencies of the test bed system for the door link. Underneath the plots four additional axes have been added. The first axis shows the groundtruth, i.e. when the subject really entered and left the room. Green triangles mark the entering into the room, while red squares denote the subject leaving the room. The next axis shows the events detected by the *Door Link Event Detection*. The third axis shows the remaining door events and their estimated type after filtering by the *Event Classification*. The type of the events is indicated by their shape and color as in the groundtruth axis. The detected subject presence is shown on the bottom axis.

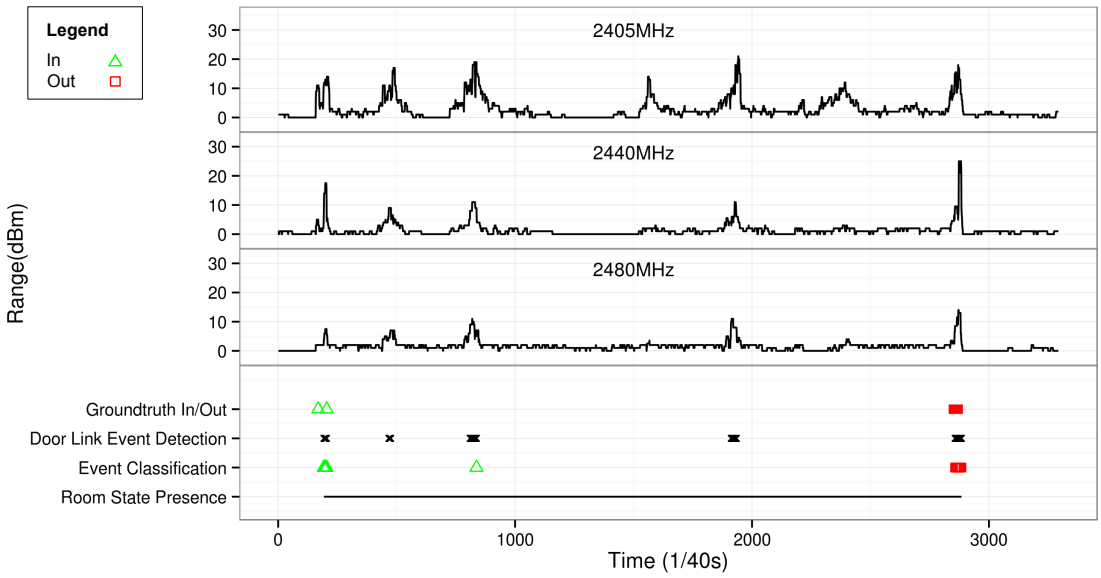



Figure 31.: Output of the LoS-Cross modules and groundtruth for a single evaluation dataset

The visualization explains the huge decrease in performance when omitting the silence link. The *Door Link Event Detection* detects many additional events of similar magnitude and duration as the entering event. The *Event Classification* discards most of these false positives and categorizes the remaining events. In conclusion the detection of entry or leave using a single link RSS alone is very challenging.

### 6.4.4 *Summary of the optimal System Configuration*

The optimal system configuration achieved a 92.86% accuracy detecting the subjects presence exactly in all evaluation data sets but one. It employed the range feature over a 0.25 s window and used all three frequencies on both links. The most important parameters was the detection window size and the silence link. Both had tremendous influence on the system accuracy (ca. 50%).

Analyzing the system performance in more detail, the reason for the failed detection becomes apparent: The calibration data set contained fluctuations much higher than what was observed for entering in the actual evaluation. Consequently,

the *Door Link Event Detection* detected no crossing event when the subject entered the room. As only a single calibration data set is used to define the threshold an outlier in this data will easily cause unsuitable thresholds. Thus, detection could be improved by requiring a second calibration run or a more sophisticated link crossing detection mechanism. As discussed later, a more practical approach may be the replacement of the offline calibration by a continuous calibration approach.

LoS-Cross uses a three stage approach to detect the presence of a subject: crossing event detection, event classification and finally room state determination. In the previous sections we have investigated the accumulated system performance using the overall Presence Detection accuracy. However, as described in Section 6.3.4 the performance depends on how well the system discriminates events into "Entering" and "Leaving" categories. For these two types of an event we consider the *Event Classification* as a binary classifier.

| Groundtruth \ LoS-Cross | Entering | Other |
|---|---|---|
| Entering | 13 | 1 |
| Other | 1 | 27 |

(a) Classification Performance for the "Entering" Event

| Groundtruth \ LoS-Cross | Leaving | Other |
|---|---|---|
| Leaving | 14 | 0 |
| Other | 0 | 28 |

(b) Classification Performance for the "Leaving" Event

Table 6.: Confusion matrices for the identification of detected events

Figure 6 shows the confusions for each type given the number of detected and undetected events over all evaluated experiments. It can be observed that the accuracy for both events is high: the "Leaving" event is correctly identified in all cases resulting in a 100% classification accuracy. However, the "Entering" event only achieves a 95% accuracy as one "Entering" event is not detected and another "Entering" events was incorrectly detected (cf. Fig. 31).

### 6.4.5 *Comparison to existing Research*

Previous RSS-based Presence Detection research also reported impressive detection results (up to 100%). For some systems, e.g. [45], these results were achieved using very elaborate evaluations (different testbeds, wider area, longer duration). Nevertheless, it is difficult to compare these results with LoS-Cross performance as the other systems really detected motion of a subject, while we infer occupancy of a specific area. This is also the reason why other systems typically profit from larger detection windows; the chance of observing subject motion increases. Hence, these systems will fail for a static subject but detect motion as soon as the subject moves again.

## 6.5 DISCUSSION

### 6.5.1 *Topology and Detection Approach*

LoS-Cross relies on a specific topology and event sequence. In contrast, previous systems, e.g. [45] use signal fluctuations on any link in the monitored area. As these systems do not require special transceiver locations, they may work given an existing wireless infrastructure. Also LoS-Cross requirements may not be ideal for some application scenarios such as intrusion detection as an intruder may not always enter through the door. However, for Presence Detection in typical Smart Home scenarios it is likely that LoS-Cross would be superior than such systems; especially when the subject performs activities with very limited motion (e.g. sleeping).

### 6.5.2 *State Machine*

In order to make LoS-Cross a robust Presence Detection system the state machine is defined in such a way that a new event belonging to the current state has no effect. That is when the event "Entering" occurs if the system is already in the state "Present" the state will not change. In Section 6.4.4 we have shown that the classification accuracy for the event determination is very high. Thus, the approach holds the potential for multi subject Presence Detection: detection of another "Entering" event could be counted as an additional subject in the room. Previously, multi subject recognition was only achieved by investigating link fluctuation frequency [95] or using a localization system with many nodes [11].

### 6.5.3 *Calibration*

LoS-Cross detection builds on a number of thresholds derived in a very brief calibration procedure which was repeated prior to every evaluation. Calibration is cumbersome and reduces user comfort. The implemented fast calibration process may also lead to suboptimal thresholds. Previous work has pointed out alternative approaches: RASID [45] uses a 2 min calibration phase before running a continuous online calibration keeping detection stable over a longer period of time. Similarly, Ichnaea [67] uses a short training period in combination with anomaly detection and particle filtering for adapting to environmental changes. LoS crossings are typically very strong fluctuations (i.e. outliers). Hence, LoS-Cross performance could be improved using outlier detection instead of calibration. In this case, if the first outlier ("Entering") is detected LoS-Cross could adjust all other thresholds accordingly. A candidate algorithm for outlier detection is the extreme studentized deviate (ESD) [65]. ESD is especially interesting as it offers the possibility to consider a priori knowledge such as the door link length directly within the algorithm.

## 6.6 CONCLUSION

In this chapter we have presented the device-free, radio-based Presence Detection system LoS-Cross. LoS-Cross uses a specific topology to detect door crossing events in a monitored area. We have evaluated LoS-Cross and achieved a 92% accuracy for exact presence detection of a single individual in an office room. We further showed the strong impact of a number of system parameters. The most important parameters being the window size for event detection and the availability of a stabilizing link. LoS-Cross demonstrates the feasibility of a topology constrained Presence Detection system. It provides more robustness than other RSS-based systems as it can recognize presence in spatial areas with walls quasi transparent to 2.4GHz radio signals. It further allows to derive presence even when the subject conducts activities with very limited motion making it an ideal building block for a holistic device-free, radio-based Activity Recognition architecture. Previous systems with comparable performance either rely on special radio signal information or employ a much larger number of nodes. Nevertheless, the requirement of a special topology makes the system less comfortable and makes it hard to transfer it to existing wireless infrastructures. On the other hand, the information that comes with the topology requirements make it likely that LoS-Cross can be realized without calibration.

With the successful evaluation of LoS-Cross we consider the requirement of a Presence Detection Inference System for a practical Activity Recognition system as described in Chapter 5 satisfied.

# 7 WiDisc: Wireless Subject Discrimination System

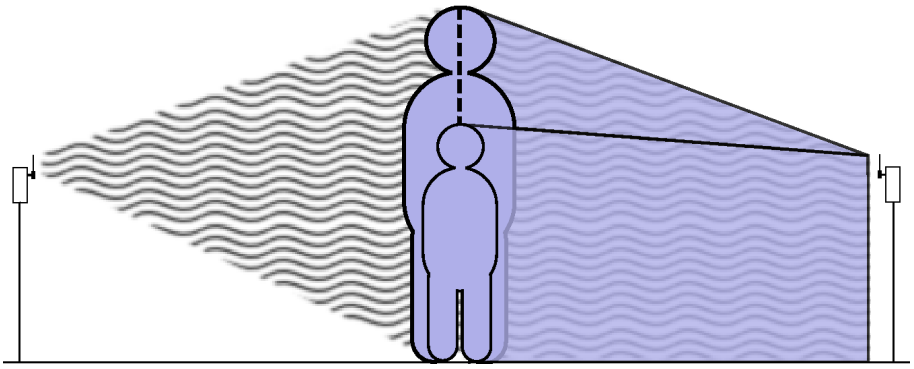
This chapter supports the validation of Thesis 3: the development of a practical device-free Activity Recognition system. In Chapter 5 we reasoned the necessity of Subject Discrimination in order to implement practical radio-based Activity Recognition. The WiDisc Subject Discrimination Inference System presents the second building block in the proposed holistic system.

This chapter has been published in our paper "Device-Free Radio-based Low Overhead Identification of Subject Classes" [71]. The structure is as follows: In the introduction we reason and illustrate the possibility of RSS-based discrimination of subjects. The next section reviews the limited related work on this topic. In section 3 we present the architecture of WiDisc. Section 4 presents the evaluation of WiDisc in which we discriminate between three different subjects. The chapter closes with a discussion and conclusion.

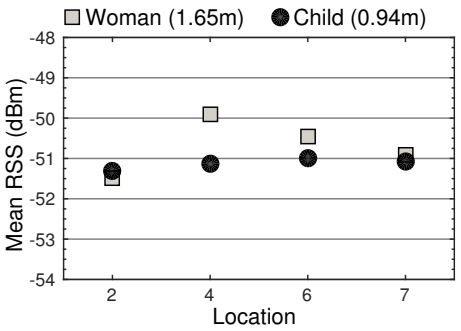
## 7.1 Introduction

Recently, researchers started to investigate how the received signal strength (RSS) can be used to differentiate between different classes of objects. In Ref. [5] and [43] the authors presented techniques to discriminate car and human. In this chapter, we take the device-free object identification a step further by differentiating between different classes of subjects. Specifically, we address the problem of differentiating between tall, medium, and small subjects. Besides the possibility to enable practical Activity Recognition subject differentiation can enable a novel set of applications including fine-grained intrusion detection forensics, parental control, among others. Additionally, subject discrimination could help to improve the accuracy of other device-free functionalities; for example, recognition algorithms may achieve better results if they consider subject specific effects. Finally, investigating radio-based subject differentiation improves our understanding of the possibilities, limitations, and especially privacy risks associated with the device-free technology.

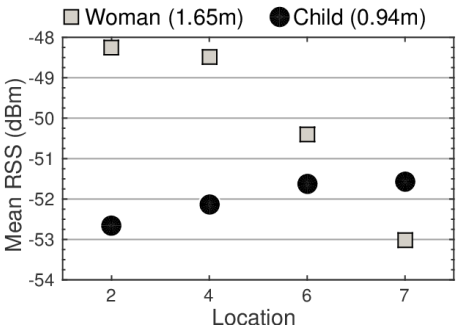
Our underlying assumption for device-free class discrimination is straight forward: objects of different sizes/dimensions alter a different number of signal paths between the transmitter and the receiver. In general, we assume that the bigger an



(a) Investigation rationale




(b) Mean RSS for woman and child on 2405 MHz at different locations.




(c) Mean RSS for woman and child on 2440 MHz at different locations.

Figure 32.: RSS differences in subjects can be observed across locations and frequencies.

object is, the more paths are affected or added by the object. To illustrate this idea, consider Figure 32a: the transmitter on the left side sends a signal to the receiver on the right. In the middle, two subjects of different heights are shown. It can be observed that the shadowed area to right of the subjects is directly influenced by their size. Figures 32b and 32c show the effect for two humans with different heights (a woman and a child) at different locations (highlighted in the testbed in Figure 35) and signal transmission frequencies. The figures show that the two classes have different RSS signatures at different locations and on different carrier frequencies. This indicates the possibility of algorithmic subject discrimination.

In this chapter, we present the Wireless Discrimination System (*WiDisc*) for device-free, radio-based subject class identification. We define a subject class as a set of subjects which affect the RSS of a link in a similar manner by their body stature or morphometrics. Thereby, a "link" or "RSS link" describes a wireless connection between a single transmitter-receiver pair independent of specific Radio Sensor settings. As Figure 32 suggests, subject classes could be differentiated using traditional fingerprinting, i.e. by comparing readings from different RSS links on single or multiple frequencies to a database with previously acquired subject class measurements. However, due to the high dependence of the RSS on the location of the subject, the required effort for such location-dependent fingerprint collection is significant. To address this issue, *WiDisc* uses 3D body

model approximation based on measurements from Microsoft Kinect sensors and electromagnetic simulation to automatically construct the fingerprint with ***zero manual training overhead***. In addition, *WiDisc* has a module to infer the most suited wireless links among the available links to further enhance accuracy and help to make on site deployment decisions.

We conducted a study in which we implemented and deployed *WiDisc* in our lab. Evaluation with three different classes at seven different locations using a wireless network of four nodes showed that *WiDisc*can accurately predict the two most important links leading to an accuracy less than the maximum achievable by only 5%. Moreover, *WiDisc* can discriminate the different classes with an accuracy of 67% accuracy ***with no manual training***. This increases to 76% if manual calibration is used, highlighting *WiDisc*s ability to trade off accuracy and training overhead.

## 7.2 RELATED WORK

In this section we review literature related to specific aspects of this chapter: the device-free discrimination of objects using radio signals and the modelling of subjects.

### 7.2.1 *Object Discrimination*

Previous device-free research has not considered the discrimination of subjects. In [73], authors explored radio-based activity recognition and reported a reduced accuracy when testing and training a device-free activity recognition system with different subjects. They proposed to investigate inter-subject differences. In Ref. [5], the authors present a system which can discriminate a car and a human. However, the significant difference in terms of speed and materials between humans and cars makes this an easier problem than our case, where objects are all human and have comparable speeds.

### 7.2.2 *Subject Modelling*

*WiDisc* depends on human body modelling to reduce the calibration overhead. In Ref. [6, 59], authors used metallic cylinders to approximate the spatial expansion of a subject. However, in order to differentiate subject classes inter-subject differences such as height and width should be emphasized, demanding more detailed modelling. The appearance of the Microsoft Kinect has fostered 3D subject reconstruction research (e.g. [37, 20, 79]). The Kinect allows spatial reconstruction at moderate cost. Current Kinect based algorithms reach good accuracies in terms of reconstruction. However, they are not openly available and need significant amounts of engineering work to be reproduced. Additionally, most of them do not allow to animate models without further processing. Thus, we consider these algorithms unavailable for our purpose. In Section 7.3.3 we describe our approach

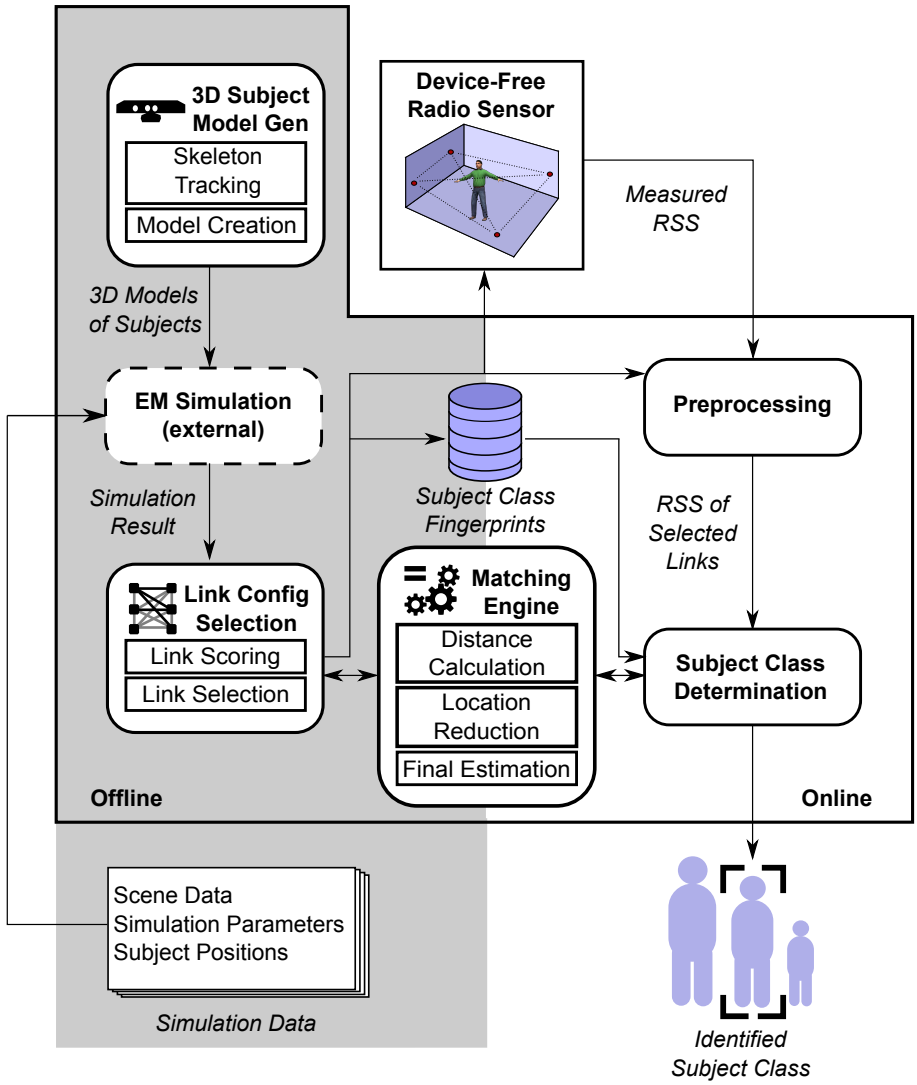



Figure 33.: WiDisc system architecture. System components placed in the shaded region are only used during the offline phase.

to automatically generate a simplified cube model that can capture more detailed representation of a human using the Microsoft Kinect sensor.

## 7.3 WIDISC SYSTEM

*WiDisc* determines the class of a subject using signal strength measurements from a wireless infrastructure. It is based on the rationale that subject classes influence RSS differently. However, location has significant influence on RSS. Hence, subject class identification across multiple locations requires collecting fingerprints from these locations from all subject classes, leading to a high calibration effort. *WiDisc* is designed to circumvent this effort. In *WiDisc*, subject classes are distinguished by comparing measurements to a fingerprint database generated through wave propagation simulation and optimization during an offline phase. During the

online phase, a distance metric is used to select the most likely class. More formally, the subject class $x \epsilon S$ is selected which satisfies:

$$x = \underset{x \epsilon S}{\operatorname{argmin}} \; d(\mathrm{SF}^x, M) \tag{22}$$

Here $S$ is the set of all subject classes, $\mathrm{SF}^x$ is the fingerprint for class $x$ over all locations, and $M$ is the real time RSS vector measured by the wireless infrastructure. $d$ is a metric function that calculates the distance between the real time measurement ($M$) and the fingerprint ($\mathrm{SF}^x$).

The section continues with an overview over the *WiDisc* architecture and is followed by a detailed explanation of each *WiDisc* module.

### 7.3.1 *Architecture Overview*

Figure 33 depicts the system architecture. *WiDisc* has two phases: the offline phase which is carried out once; and the online phase in which continuous recognition is performed. The modules active during the offline phase are placed in a shaded rectangle in the figure.

In the offline phase, the 3D body model is constructed using the *3D Subject Model Generation*. Based on this, the electromagnetic (EM) simulation is conducted using the *EM Simulation* module to generate the simulation fingerprints for different subject classes at different locations. Finally, the most relevant links are identified using the *Link Configuration Selection* and the *Matching Engine*. In offline mode, the *Matching Engine* computes the distance between each simulation RSS vector and all other simulation database vectors for each possible link set. The output of the offline phase is a simulation fingerprint database containing only those links estimated to be most relevant by the *Link Configuration Selection* module. For this reason, *WiDisc* can also be used to create suggestions for the installation of a new wireless infrastructure to reduce the hardware cost as well as remove noisy links.

During the online phase, the *Radio Sensor* acquires a measurement which is processed by the *Preprocessing* for noise reduction. The processed measurement serves as input to the *Subject Class Determination* module, which uses the simulation fingerprint database and the *Matching Engine* to decide on the subject class of a measurement. As shown in the introduction the use of multiple frequencies can be beneficial to distinguish subject classes from each other. Therefore, *WiDisc* has been designed to specifically consider multi frequency measurements of links.

In the remainder of this section, we give the details of the various *WiDisc* modules.

### 7.3.2 *Device-Free Radio Sensor*

A radio sensor is a configuration of wireless transceivers installed in the area of interest (cf. Chapter 4). It acquires the RSS between all transceivers and relays them to the *WiDisc* system. The radio sensor typically has a number of parameters

such as carrier frequency, number of transceivers, and installation locations. The *WiDisc* offline mode can assist in selecting a suited Radio Sensor configuration by repeatedly evaluating the results of the *EM Simulation* and *Link Configuration Selection* modules, given different parameter settings. Alternatively, the sensors may already be installed, e.g. WiFi infrastructure or Smart Home installations using IEEE 802.15.4 radios. In this case, the information supplied by the *Link Configuration Selection* module is utilized in the *Preprocessing* module to remove noisy links and hence provide better subject class discrimination performance.

### 7.3.3 *3D Subject Model Generation*

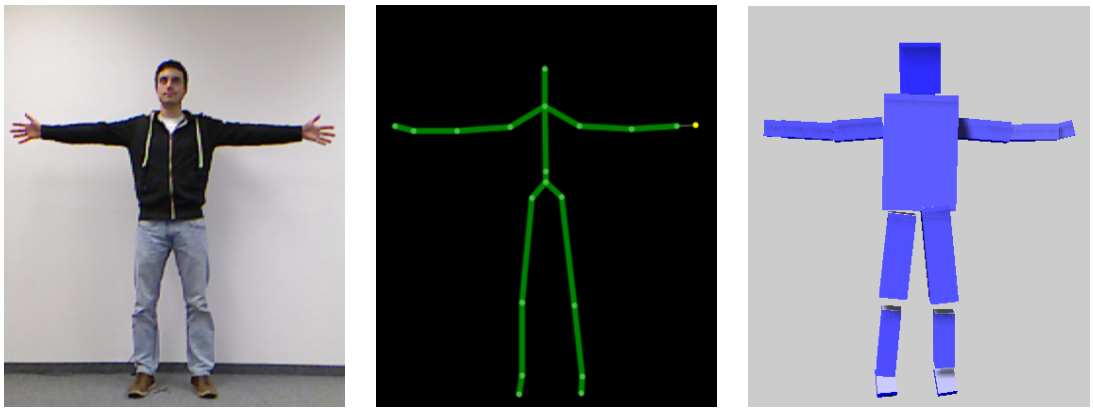
This module creates a simple 3D box model approximation of a subject. It consists of two submodules: *Skeleton Tracking* and *Model Creation*. *Skeleton Tracking* employs the Microsoft Kinect sensor to create a 2D skeleton of the subject. Using this skeleton, the *Model Creation* submodule then constructs a 3D box model. Despite its simplicity, it offers much more details than the simple cylinder body model typically used in the device-free literature, e.g. in [59, 6].

#### *Skeleton Tracking*

To construct the 3D model, the subject stands facing a Kinect sensor in good operating distance. The subject pose should be chosen so that the Kinect skeleton tracker can estimate position and rotation of all joints. An example of a good pose is the T-pose shown in Figure 34a. We found that a good operating distance is between 1.80m and 3.80m, if the Kinect is mounted in a height of 0.9m. While the subject stands in front of the Kinect, the module uses the skeleton tracker of the Kinect Windows SDK to derive parameters for the dimensions of the boxes resembling the human model. The tracker describes a subject by estimating the position and rotation of the 19 major joints and the corresponding bones.

#### *Model Creation*

Given the spatial position of the joints, we derive the position and height (length along the y-axis) of 15 3D boxes approximating the spatial expansion of a subject's limbs and torso. After the dimensions of the boxes and initial positions resembling the subject model are computed, the subject is continuously tracked. During tracking, the box model follows the motions of the subject. To acquire our subject model for the experiment the subject was instructed to take the pose "standing still". The 3D box model of this pose was exported for simulation. Figure 34 shows an optimal scanning pose, the skeleton tracker output, and the reconstructed 3D model. The subject material is set based on previous research [6].



(a) Subject scanning T-pose (b) Skeleton tracker result (c) Constructed 3D cube model

Figure 34.: 3D cube model construction

| | |
|---|---|
| Simulation mode | Phase-coherent ray tracing<br>Transmission and reflection: Fresnel coefficients<br>Diffraction: uniform theory of diffraction |
| Effects | Transmission, reflection, diffraction |
| Path number | Unlimited (maximal path loss: 200dB) |
| Interactions per path | 4 transmissions<br>2 reflections<br>2 diffractions<br>3 reflections and diffractions |

Table 7.: Simulation settings.

### 7.3.4 *Electromagnetic Simulation Module*

This module predicts the subject specific RSS. Simulation is performed by interfacing an external wave propagation simulator. The input data to the simulator were the room dimension, materials, furniture, and locations of the subjects in the room. We also provided the simulator with the 3D subject models constructed by the *3D Subject Model Construction* module and the transceiver positions and carrier frequencies. The simulation was performed using the settings in Table 7. The simulation run time per location and subject was ~4 h on a mid 2014 desktop PC with i7 3770K processor and 16GB RAM.

Simulation was performed for each subject and position, resulting in a predicted RSS fingerprint matrix $SF^x$ per subject class $x$. Each matrix had the same dimension $m \times n$. Thereby $m$ is the number of positions $|P|$ and $n$ is the number of links $|L|$ with the following structure:

$$\mathrm{SF}^x = \begin{pmatrix} RSS_{p_0,l_0} & \cdots & RSS_{p_0,l_{n-1}} \\ \vdots & \ddots & \vdots \\ RSS_{p_{m-1},l_0} & \cdots & RSS_{p_{m-1},l_{n-1}} \end{pmatrix} \tag{23}$$

A cell in this matrix $RSS_{p_i,l_j}$ is a vector of simulation-based predicted RSS of all used carrier frequencies for a 3D model of subject class $x$ positioned at location $p_i \epsilon P$ on one of the simulated links $l_j \epsilon L$.

### 7.3.5 *Link Configuration Selection*

This module analyzes the simulation results to estimate the set of links giving the best results for subject class discrimination. This is useful for optimizing the Radio Sensor configuration for a new deployment and for eliminating noisy links. Its input is the simulation fingerprints of all subject classes. The output is an optimized simulation fingerprint database containing only the estimated best links. The relevant links discovered by this module are also used during the online phase in the *Preprocessing* module.

The module processes simulation results in two steps: First, the *Link Scoring* submodule scores each link with its carrier frequencies based on cross validation performance on the simulation database for all link combinations. In the second step, the *Link Selection* submodule analyzes the link scores to select the best link configuration.

#### *Link Scoring*

This submodule scores each simulated link by its subject class discrimination accuracy. The accuracy is computed using cross-validation, where the simulation fingerprint is split into training and testing over all link combinations. The number of folds is taken as all possible combinations of subjects, positions, and links. For each link configuration, the accuracy is calculated using the *Matching Engine* for different matching functions (described in Section 7.3.7).

In addition, an optional variant, i.e. aggregating all frequencies of a single physical link, is also evaluated as compared to using each frequency independently.

The intuition is that combining all link frequencies may balance simulation errors in single frequencies but conserve general trends. The aggregated fingerprint is obtained by averaging the fingerprints over all frequencies of a given physical link.

#### *Link Selection*

The goal of this module is to select the combination of links that gives the best score taking into account the possible ties in score between different configurations. First, the configuration with the best score calculated by the *Link Scoring*

submodule is selected. If there are ties between a number of configurations, the configuration with the least number of links is selected (to optimize the hardware used). If there are still ties, the configuration with the best performing links over all configurations is selected. To determine the best performing links, we assign an overall to each link based on the achieved accuracies for each configuration it appeared in, weighted by the number of other links in this configuration. More formally, the overall link score for a specific link $l$ is calculated as:

$$\text{Score}(l) = \sum_{c \in C} \frac{Accuracy(c)}{|c|} \tag{24}$$

Where $C$ is the set of all configurations that include link $l$, $Accuracy(c)$ is the achieved accuracy for configuration $c$, and $|c|$ is the number of links in configuration $c$.

The fingerprints of the best selected links are stored in the subject class database for use during the online phase.

### 7.3.6 *Preprocessing*

The *Preprocessing* module prepares the measured RSS from the *Radio Sensor* for the *Subject Class Determination* module. It first averages the measured RSS over a time window (five seconds in our experiments) to remove fluctuations. In frequency aggregation mode (cf. Sec. 7.3.5), it also averages the RSS over all the frequencies of a single link. Finally, the module selects the best links chosen by the *Link Configuration Selection* module in the offline phase.

### 7.3.7 *Matching Engine*

This module selects the best subject class for a given input measurement $M$ based on the simulation fingerprints constructed during the offline phase ($\text{SF}^x$). The core of this module is a matching function $d(\text{SF}^x, M)$ that determines the best subject class. Recall that each fingerprint for a subject class is a matrix. Each cell contains a simulated RSS value. The rows and columns represent the different fingerprint locations and links, respectively. For simplicity of presentation, and without loss of generality, we assume a single frequency carrier. To obtain the best subject class we use three steps: distance calculation, location reduction, and final estimation. ***We experimented with different matching functions for each of the three steps.*** Here we list only the functions that performed best.

#### *Distance Calculation*

The first step is to calculate the $L_1$ norm between each row in the fingerprint ($\text{SF}^x$) and the online measurement vector $M = (M_{l_0}, ..., M_{l_{n-1}})$. The output of this step is a distance matrix $\text{dSF}^x$ for each subject class representing the difference between

the measurement vector and the RSS vector at each location in the fingerprint for a specific subject:

$$\mathrm{dSF}^x = \begin{pmatrix} |RSS_{p_0,l_0} - M_{l_0}| & \cdots & |RSS_{p_0,l_{n-1}} - M_{l_{n-1}}| \\ \vdots & \ddots & \vdots \\ |RSS_{p_{m-1},l_0} - M_{l_0}| & \cdots & |RSS_{p_{m-1},l_{n-1}-M_{l_{n-1}}}| \end{pmatrix} \tag{25}$$

*Location Reduction*

The next step is to reduce the $\mathrm{dSF}^x$ matrix into a single row vector ($\mathrm{lSF}^x$) by projecting it into the links dimension. This is achieved by selecting from each column the entry with the minimum value (representing the RSS distance of the best candidate location for this link):

$$\mathrm{lSF}^x = \left( \min \begin{bmatrix} \mathrm{dSF}^x_{0,0} \\ \vdots \\ \mathrm{dSF}^x_{m-1,0} \end{bmatrix}, \ldots, \min \begin{bmatrix} \mathrm{dSF}^x_{0,n-1} \\ \vdots \\ \mathrm{dSF}^x_{m-1,n-1} \end{bmatrix} \right) \tag{26}$$

*Final Estimation*

The final step is to select the best candidate subject class based on the projection vectors $lSF^x$ for each class $x$. For this purpose, we define three candidate matching functions ($\mathrm{mf}_1$, $\mathrm{mf}_2$, $\mathrm{mf}_3$) that estimate the subject class by operating on a single matrix D. This matrix is a concatenation of the projection vectors. Hence, it has a row for each row vector $\mathrm{lSF}^x$. The dimension of D is "number of subjects" $\times$ "number of links" ($|S| \times n$):

$$\mathrm{D} = \begin{pmatrix} \mathrm{lSF}^0 \\ \vdots \\ \mathrm{lSF}^{|S|-1} \end{pmatrix} \tag{27}$$

**Matching Function 1 ($\mathbf{mf}_1$):** returns the subject class having the global minimum over all subjects and links in the D matrix:

$$\mathrm{mf}_1(\mathrm{D}) = \underset{x \in S, i=0..n-1}{\operatorname{argmin}} \mathrm{D}_{x,i} \tag{28}$$

**Matching Function 2 ($\mathbf{mf}_2$):** The idea of $\mathrm{mf}_2$ is that the correct class will have the smallest sum of RSS differences among all links. That is, it will lead to the closest match on all links:

$$\mathrm{mf}_2(\mathrm{D}) = \underset{x \in S}{\operatorname{argmin}} \sum_{i=0}^{n-1} \mathrm{D}_{x,i} \tag{29}$$

**Matching Function 3 ($\mathrm{mf}_3$):** is a majority voting function. In particular, it returns the subject class that is selected by the majority of the links as the best class:

$$\mathrm{mf}_3(\mathrm{D}) = \underset{x \in S}{\operatorname{argmax}} |Y| : \forall y \in Y \leftrightarrow y \in \mathrm{G}(\mathrm{D}) : y = x \tag{30}$$

G is a row vector whose entries represent the best subject class selected for each link:

$$\mathrm{G}(\mathrm{D}) = \left( \underset{x \in S}{\operatorname{argmin}} \begin{bmatrix} \mathrm{D}_{0,0} \\ \vdots \\ \mathrm{D}_{|S|-1,0} \end{bmatrix}, \ldots, \underset{x \in S}{\operatorname{argmin}} \begin{bmatrix} \mathrm{D}_{0,n-1} \\ \vdots \\ \mathrm{D}_{|S|-1,n-1} \end{bmatrix} \right) \tag{31}$$

## 7.4 EXPERIMENTAL EVALUATION

In this section, we provide the performance evaluation of *WiDisc*. We start by describing our experimental testbed followed by analyzing the effect of the different parameters. The section is concluded by the detailed investigation of the performance of the best parameter configuration. The default parameter values are given in Table 8.

| **Parameter** | **Default Value** |
|---|---|
| Frequency Aggregation | Mean |
| Matching Function | $\mathrm{mf}_1$ |

Table 8.: Default values for the *WiDisc* parameters.

### 7.4.1 *Testbed and Experiment Execution*

*Setup*

The testbed consisted of four 2.4 GHz NXP Jennic 5139 IEEE 802.15.4 transceivers with omnidirectional SMA Gigaant swivel antennas (cf. Appendix A). Transceivers measured the RSS on six links and three different carrier frequencies ($f_c$ = 2405 MHz, 2440 MHz, 2480 MHz) with 40 Hz. The testbed was installed in an office room of 4.0x5.33 m size and 2.70 m height (Figure 35). The room had lightweight partition walls, four windows, and a concrete ceiling. It was equipped with pressboard furniture, a single office chair, and two desktop computers. The transceivers were placed on the left and right side of the door, on the window front, and in the middle of one side wall at 0.80m height.

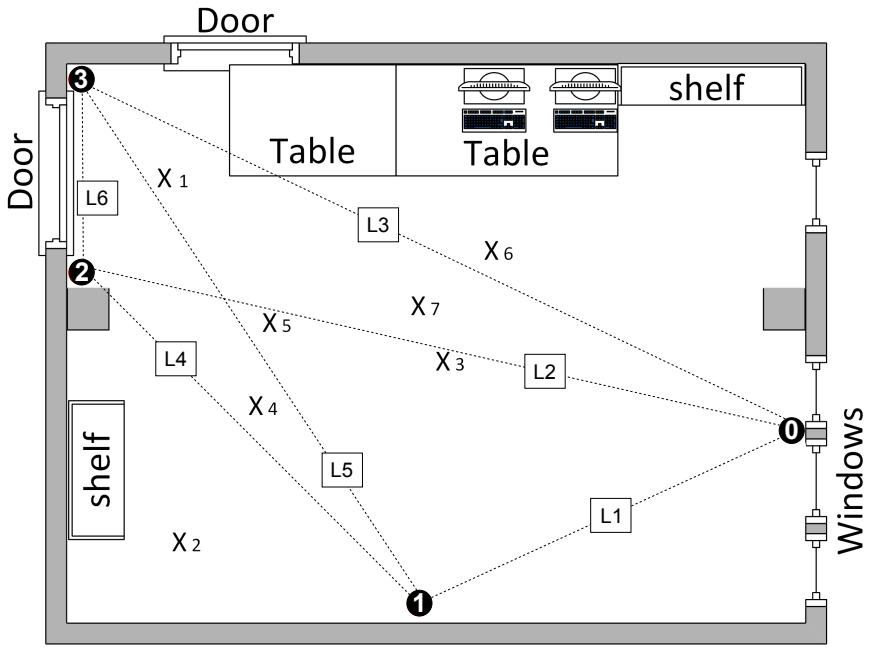



Figure 35.: Room floor plan with transceiver locations (black circles), subject positions (X's) and links (rectangles).

For the experiment we selected three different subject classes each represented by a subject as follows:

- Small subject class: child, 0.94 m, 13.5 kg
- Medium subject class: woman, 1.65 m, 60 kg
- Tall subject class: man, 1.95 m, 79 kg

*Offline Phase Preparations*

Each of the subjects was scanned with a single Kinect camera and a 3D box model was created using the *3D Subject Model Construction* module described in Section 7.3.3. A room model was created with furniture, doors and windows and transceivers. For each of the seven test locations of the three subjects, a 3D scene file was generated in which the subjects were located at one of the positions facing the window side. The subjects' pose was standing still with arms hanging loosely at the sides. A total of 21 (3×7) simulations were conducted. The result of the simulation were three simulation fingerprint matrices with dimensions of 7×6 ($n = 7$ and $m = 6$), each corresponding to a single subject class. That is, $SF^x$ with $x \in \{\text{Small, Medium, Tall}\}$. Note that each cell in this matrix is a three element vector of the RSS on the three carrier frequencies. Figure 36 shows a 3D visualization of a simulation scene including the "Tall" 3D model. In the figure, white rectangles indicate the position of the transceivers. Using the simulation fingerprint for each subject ($SF^x$), the *Link Configuration Selection* submodule was used to construct an estimation of the best link count and best link set for all

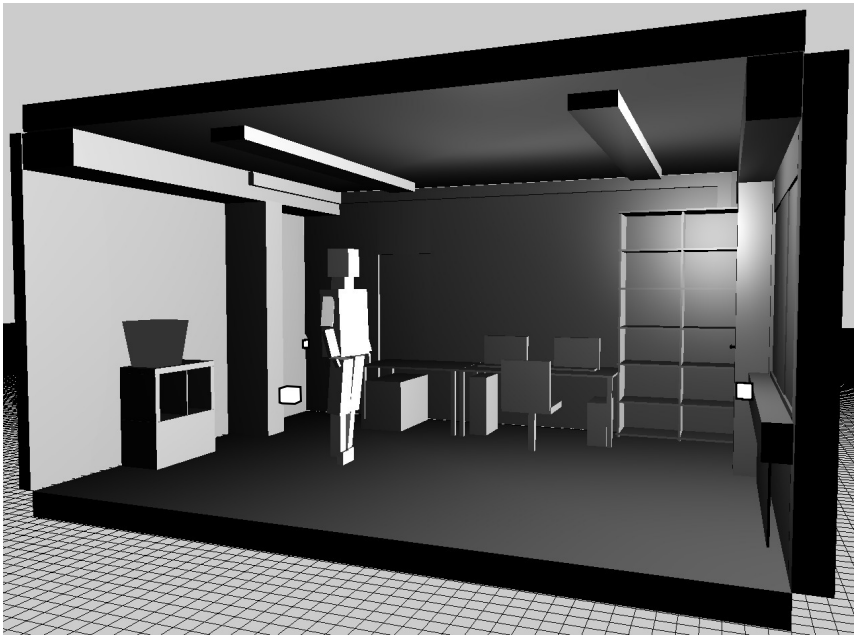


Figure 36.: 3D room model with box model of tall subject at $X_4$ and three visible transceiver positions (white rectangles), position of transceiver 2 is occluded by the column left of the door.

selected configuration combinations (cf. Sec. 7.3.5). Finally, the optimized/reduced fingerprint matrices with the selected links were stored in the fingerprint database. Note that no manual fingerprinting was performed/needed for *WiDisc*.

*Online Phase*

Each subject stood still for 5 s at each of the seven positions. Figure 37 shows the tall subject during the measurement in the testbed with transceiver 3 highlighted in the far right corner of the room. Concurrently and continuously, the RSS measurements were produced by the *Radio Sensor* and relayed to the *Preprocessing* module running on a notebook PC connected to one transceiver via USB. There, the data was averaged over 5 s to a single vector and, if aggregation was enabled, the different frequencies were averaged. The measurements were then forwarded to the *Subject Class Determination* module, which performed classification through the *Matching Engine*.

### 7.4.2 *System Performance for different Parameter Settings*

In this section, we explore the performance of *WiDisc* for different configuration settings in the *Link Configuration Selection* and the *Matching Engine*, namely the effects of the frequency aggregation (cf. Sec. 7.3.5) and different matching functions (cf. Sec. 7.3.7). For each configuration, we show the ***estimated performance***, i.e. performance estimated by the *Link Configuration Selection* module by cross-validation, and the ***actual performance*** achieved in the experiment. The two

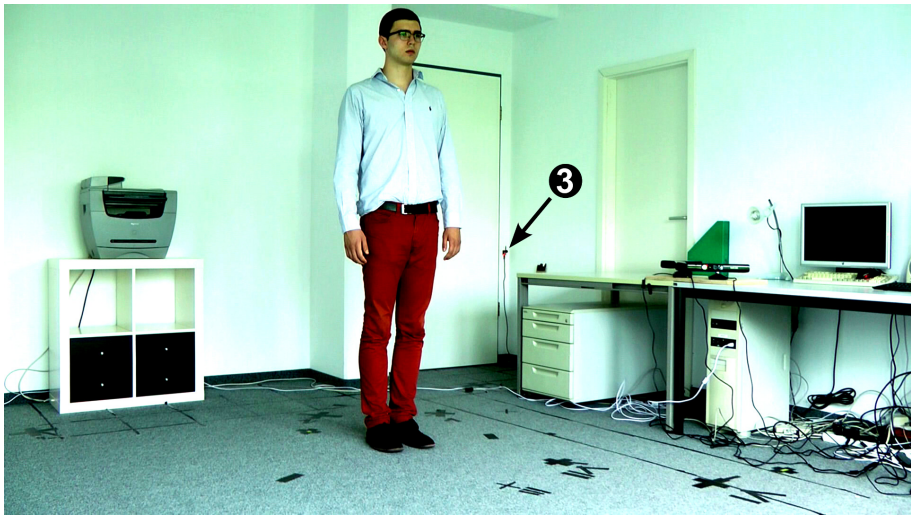



Figure 37.: Tall subject at location $X_4$ with transceiver 3 in the background.

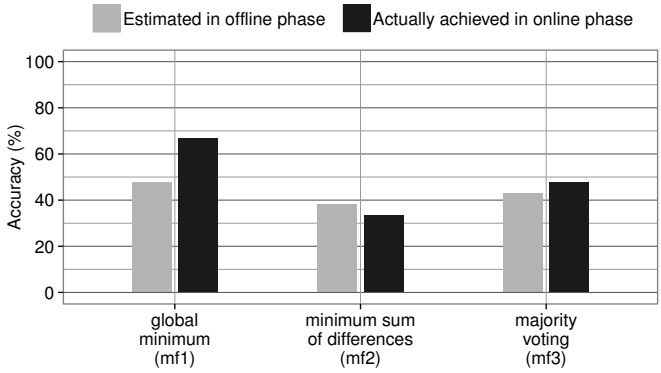



Figure 38.: The effect of different matching functions on the performance.

metrics show how closely the simulation matches the reality. Except for the respective parameter being tested, all other parameters were set to default values (Tab. 8).

*Effect of different Matching Functions*

Figure 38 shows the estimated and actually achieved accuracies for the three matching functions defined in Section 7.3.7, namely the global minimum ($mf_1$), minimum sum of differences ($mf_2$), and majority voting ($mf_3$). The best performing matching function was the global minimum matching function $mf_1$, with an actual classification performance of 67%. Matching function $mf_2$, the minimum sum of differences, performed worst. The success of the actual class prediction during the online phase depends on the correctness of the simulation conducted in the offline phase. The superior performance of the global minimum matching function, which selects the subject class with the smallest overall distance, highlights the

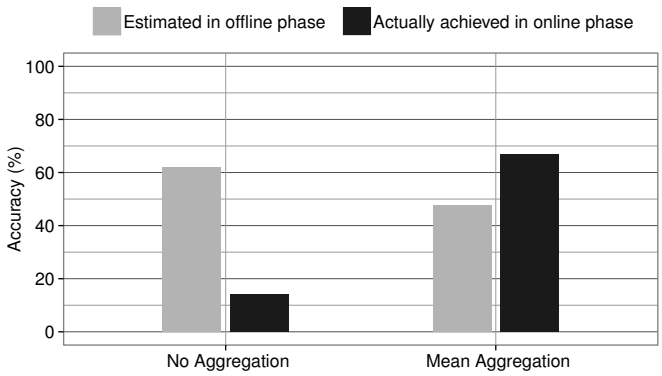



Figure 39.: The effect of frequency aggregation on the performance.

observation that the RSS is highly location dependent. Thus, we cannot fuse the scores at different locations to improve accuracy as attempted in $mf_2$ and $mf_3$.

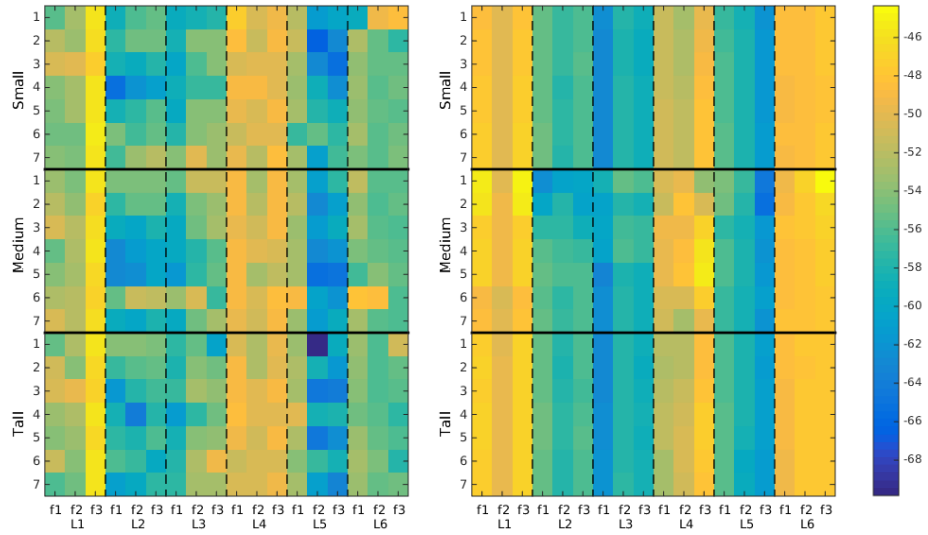
*Effect of Aggregation*

Here, we investigate how much accuracy is gained using the aggregation of the RSS from the different frequencies as described in Section 7.3.5. The effect of this parameter is depicted in Figure 39. The figure shows that link aggregation significantly enhances the actual achieved accuracy from 14% to 67%. More interestingly, the predicted accuracy through simulation was negatively correlated with the actual achieved accuracy.

To investigate this in more detail, Figure 40 shows a heat map of simulation data from the offline phase and the measurement data from the actual experiment in the online phase. We observe that there is a significant difference between both heat maps. Specifically, the measurement shows a smoother RSS in respect to subjects and locations compared to simulations. Frequency averaging acts as a smoothing filter, making simulation data more similar to the measurements (cf. Fig. 41), and hence achieving better classification results.

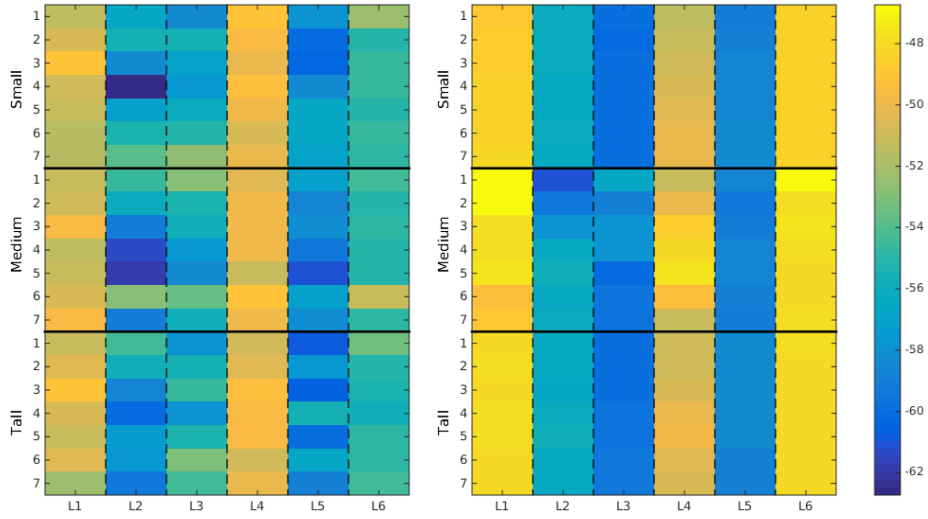
*Overall Performance*

DETAILED SYSTEM PERFORMANCE FOR THE BEST CONFIGURATION The best performance was achieved when employing frequency mean aggregation and using the global minimum matching function $mf_1$. The detailed performance of this configuration is shown in Figure 42. The figure shows that the estimated accuracy on simulation data during the offline phase was 48% (using links ***L5 and L4*** predicted by the *Link Configuration Selection* module), the actually achieved performance in the experiment during the online phase was 67% (using the same two links: ***L5 and L4***), and the best possible achievable accuracy was 72% (obtained when using the three links: ***L5, L4, and L2***). The difference between the actual achievable accuracy and the best possible accuracy was due to the different links used ***(L5 and L4)*** only compared to ***(L5, L4, and L2)***. The latter configuration, the




(a) Simulation (b) Measurement

Figure 40.: RSS of simulation and measurement for all three subject classes and different positions ($X_{1..7}$) over all links (L1..L6) and frequencies (f1, f2, f3).




(a) Simulation (b) Measurement

Figure 41.: RSS of simulation and measurement for all three subject classes and different positions ($X_{1..7}$) over all links (L1..L6) **averaged over frequencies**.

one that gives the best configuration, cannot of course be known in advance during the offline phase and the *Link Configuration Selection* module tries to estimate it. The 5% difference between the achieved performance and the best possible accuracy highlights the good selection criteria of the *Link Configuration Selection* module. Figure 43 shows the estimated and actual scores and links ranking for this particular configuration. The figure further highlights that the module can

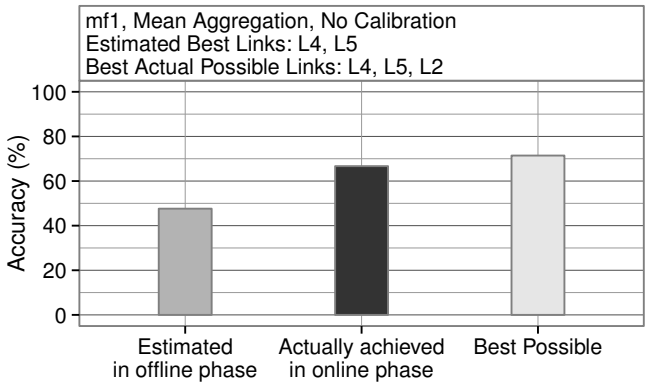



Figure 42.: Performance of the best configuration.

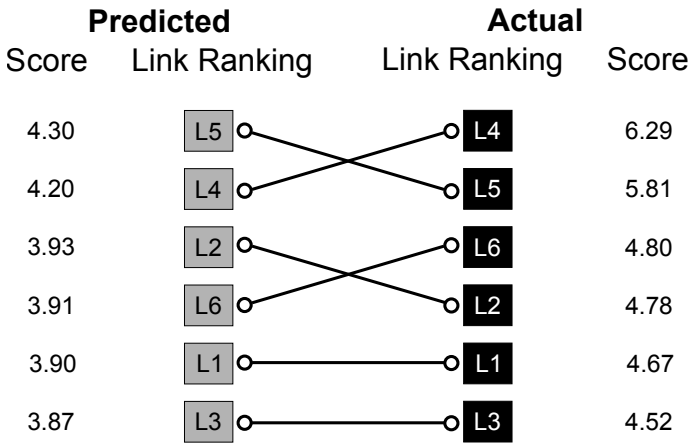



Figure 43.: Estimated and actual link scores and ranks for the best configuration.

accurately rank the different links, with at most one error in the rank of the links, leading to accurately estimating the best link(s) as described in Section 7.3.5.

SUBJECT CLASS CONFUSION MATRIX Finally, we investigate the confusion matrix of the different subject classes for the best configuration (Tab. 9) The "Tall" subject class had the highest true positive rate (85%), the "Medium" subject class had the second highest rate (71%), and the "Small" subject class had the lowest true positive rate (42%). The small and medium subject classes were confused with each other more frequently. One reason for the lower accuracy in the case of the "Small" subject class is the height of the installed nodes, which barely cut

| Actual \ Predicted | Small | Medium | Tall |
|---|---|---|---|
| Small | 3 | 3 | 1 |
| Medium | 2 | 5 | 0 |
| Tall | 0 | 1 | 6 |

Table 9.: Confusion matrix for the subject class discrimination using *WiDisc*

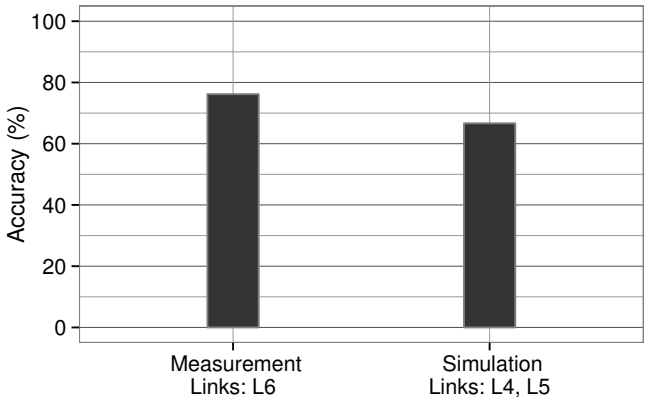



Figure 44.: Results for manual fingerprinting compared to simulation.

the head of the "Small" subject class. This led to a reduced effect on the received signal and hence lower classification accuracy.

### 7.4.3 *Comparison with Manual Fingerprinting*

Figure 44 compares the accuracy of *WiDisc* to manual fingerprinting. For the manual fingerprinting, the subject stood for 5 seconds at each location to create the training database. The figure shows that manual calibration can lead to an accuracy of 76%, compared to 67% using the automatic fingerprint employed by *WiDisc*. This slight decrease in accuracy is motivated by the significant decrease in the calibration overhead required for manual calibration.

## 7.5 DISCUSSION

### 7.5.1 *Simulation Model Complexity*

A requirement of our approach are 3D models of the environment and subjects. Using the Kinect sensor to automatically construct the human body model helps in reducing this overhead. In addition, Smart Home design companies may have the room models readily available. Further on, current projects such as Google's project Tango promise to bring Kinect-like scanning abilities to normal smart phones, opening up new possibilities for acquiring not only room data, but also object models. At the same time current research already provides the means to automatically generate floor plans based on inertial and device-free sensing [7].

### 7.5.2 *Subject Classes*

The initial rationale behind *WiDisc* is that subject class differences such as height are reflected in RSS and may be leveraged for discrimination. While this is supported by our evaluation, it remains fuzzy what exactly determines a subject class.

This should be refined in future work by investigating a larger number of subjects and considering additional parameters.

### 7.5.3 *Combined Tracking and Subject Class Differentiation*

This chapter focused on differentiating different subject classes. However, since we already have a location-based fingerprint, it may be possible to concurrently track the subject. Although there is a large cropus of research in the device-free literature for subject tracking, e.g. [42, 66, 92], the combination with subject class identification provides opportunities for further improvements.

### 7.5.4 *Robustness*

Strong environmental changes such as furniture changes or additional subjects will likely void the generated fingerprints. Computing fingerprints for all possible situations in advance may be feasible in certain scenerios. Another potential solution is to focus on subject-specific signal variations in suited locations. For instance, when crossing doors, at least in a single subject case this can be detected with good accuracy using CSI [90]. In addition, signal variations are more robust to environmental changes [45]. Thus, one goal of future work should be to investigate if incurred RSS fluctuations also depend on subject parameters. However, following the *WiDisc* approach may then be more challenging as the model must be correctly animated. Nevertheless, using 3D modelling, tracking and simulation is beneficial as it helps to correlate subject action and sensor data with greater precision than in classical groundtruth annotation. Clearly, *WiDisc* performance strongly depends on the correctness of the simulation. Our results were achieved using maximal simulation settings without calibration. Inherent hardware noise and simulation errors were compensated by low-pass filtering over a high number of measurements (200) and frequency aggregation. In other settings parameterization of simulation given transceiver characteristics may be necessary. Alternatively, the system may be calibrated in-situ, optimizing simulation parameters to fit measurements.

## 7.6 CONCLUSION

In this chapter, we presented the *WiDisc* subject class discrimination system based on device-free, radio-based sensing. *WiDisc* uses 3D subject model generation and electromagnetic simulation to reduce the calibration overhead. Moreover, it provides an algorithm for predicting the best wireless links to differentiate subject classes in an existing or planned wireless infrastructure. Experimental evaluation using an actual deployment showed that *WiDisc* can provide a 67% accuracy, missing the accuracy of the best possible configuration by only 5%. Moreover, *WiDisc* is only 9% less accurate than a system based on manual calibration with significant reduction in overhead. We also relayed some insights gained from

designing and evaluating the system. The system may be further expanded in different directions including experimenting with other matching functions, combining subject identification with other device-free functionalities, implementing new applications based on *WiDisc*, among others.

With the successful evaluation of *WiDisc* we consider the requirement of a Subject Discrimination Inference System for a practical Activity Recognition system as described in Chapter 5 as being satisfied.

# 8 RFHAR: Device-free, Radio-based Human Activity Recognition

This chapter supports the validation of Thesis 3: the development of a practical device-free Activity Recognition system for which a holistic architecture was proposed in Chapter 5. In this chapter we develop and evaluate the radio frequency-based Human Activity Recognition Inference System (RFHAR). It is the final component of the Inference System Ensemble in the proposed holistic architecture.

The chapter is structured as follows: The introduction provides a brief overview over the methods and results used in this chapter. The next section reviews related work. The following section presents the architecture of RFHAR. In the next section the four major measurement campaigns for RFHAR optimization and evaluation are described. Optimization and evaluation results are provided in two following sections. The chapter is closed with a discussion and a conclusion.

## 8.1 Introduction

In this chapter we present RFHAR. RFHAR is a purely received signal strength based Activity Recognition system. Previous work has either conducted feasibility studies using cross-validation (cf. Chapter 3), employed rarely available radio signal characteristics [90] or focused on gestures conducted in close proximity to a device [2].

To close this gap we present RFHAR. RFHAR is an RSS-based human activity recognition system which determines user activities. It was designed to be a robust, online recognition system which can operate in a realistic environment with only few transceivers (n=4). To achieve this goal, the RSS characteristics of human activities were studied prior to designing the system. We identified three types of activity categories. Of each category we selected an exemplary activity: "Standing", "Waving" and "Walking" for RFHAR optimization and evaluation. We developed a special optimization and evaluation strategy. Optimization sought to identify the best configurations of the RFHAR system while evaluation defined various challenging tests of practical relevance not considered previously. Hence, the contribution is three-fold:

- Categorization of human activities based on their impact on the Radio Sensor
- Implementation of the optimized RFHAR system
- Evaluation regarding aspects of practicability and robustness

Optimization and Evaluation were conducted using four different measurement campaigns which are recorded in two different testbeds spanning a time frame of more than a year. During optimization we showed that a novel frequency fusion approach and a new feature improved recognition accuracy considerably. Using independent training and test sets we showed that accuracies of 95% and above can be reached when considering only three activity locations. Considering nearly all possible locations in the room resulted in an accuracy between 80-95%. We showed that accuracy stayed in these boundaries even when the amount of training data was reduced from 8 min down to 1 min or when training data is up to 10 days old. We further demonstrated that a minor change in the Radio Sensor installation reduced median accuracy from 89% to 84% but still allowed to differentiate between "Walking" and "other activity". If only considering the activities "Walking" and "Standing" the achieved accuracy typically was 95% and above even when the state of the room door was changed.

## 8.2 RELATED WORK

### 8.2.1 *Activity and Gesture Recognition using Software-Defined Radios*

In Scholz et al. [74] we presented the first online system to classify a set of three contexts using a single link between two Software-Defined Radios (SDR). The contexts were walking, talking on the telephone and the state of the door of the room. For each of these contexts a single threshold-based classifier was trained and evaluated. Over 10 experiments in which the system was repeatedly trained and tested an average accuracy of 88% was achieved directly after training.

In Ref. [82] we conducted a two fold study. First, we investigated the recognition accuracy for a system employing sender and receiver (active Radio Sensor). Second, we investigated the performance of a receiver only setup in the same setting (passive Radio Sensor). In the first study, a sender and a receiver SDR were installed in a distance of 2 m. We conducted four activities ("Crawling", "Walking", "Lying" and "Standing") in 11 locations. We also recorded data when the space was empty. Locations were distributed around the receiver. Hence some locations were in the LoS while others weren't. Using 10-fold cross validation we achieved a 72% accuracy using the full dataset. Training and testing the Activity Recognition only on single locations led to an increased recognition performance of 80% on average. The best performance was reached for activities which were closer to the receiver or close to LoS. Accuracies achieved using the passive Radio Sensor were only comparable (70%) to the active Radio Sensor when the subject was in close proximity to the receiver (0.5 m). However, results should be interpreted with care as they were achieved using cross-validation.

In Ref. [77] and [78], Shi et al. reported on improved results over the aforementioned study. They employed a single SDR to receive FM radio station signals in order to recognize activities from three different subjects. They used a two-stage classification approach for classifying a total of six activities which were conducted in a 2x2 m area around the SDR. In the first phase the activity type is discriminated, i.e. dynamic vs static activity. In the second stage the activities are identified. They recorded a 3 min sample of each activity and use 5-fold cross validation with the classifiers SVM, k-NN and decision tree for feature subset selection. They found that the two-stage classifier (best result: SVM 87% accuracy) showed slightly improved results compared to a single stage classifier (best result k-NN 84% accuracy). They did not report recognition differences across subjects. Again results should be interpreted carefully due to the use of cross-validation.

Pu et al. [61] presented WiSee a gesture recognition system leveraging SDR for recognizing a set of gestures. WiSee can classify 9 whole body gestures and achieved an average accuracy of 94%. WiSee leverages Doppler shifts extracted from OFDM transmissions from WiFi. Pu et al. also showed that their system is even functional in non line of sight and through the wall scenarios.

### 8.2.2 *Activity Recognition using Channel State Information*

Wang et al. [90] presented the E-Eyes activity "identification" system which can discriminate eight activities and different walking routes. E-Eyes relies on channel state information (CSI) available in some 802.11n hardware. The system utilizes three links, the sliding variance, the earth movers distance and the fact that some activities typically occur in specific locations. In the experiment, CSI was sampled with 20 Hz. As previously discussed (cf. Chapter 6.2.3) activities which could not be discriminated are identified using a coarse location determination which is based on doorway crossing detection. They conducted experiments repeatedly in two testbeds over four months with four subjects. The system was trained on a single day and tested on a not closer specified number of following days. The authors reported an impressive accuracy of up to 97% when using three links and 90% when using a single link for both testbeds. Nevertheless, the authors noted that activity profiles for in-place activities cannot be easily transfered to different locations and therefore training for each specific location is required.

### 8.2.3 *Activity and Gesture Recognition using Received Signal Strength*

In Chapter 3 we showed the general feasibility of device-free Activity Recognition using RSS. We outfitted a standard office room with a total of eight transceivers in which six different activities and the empty room were recorded. Using only four transceivers a performance of up to 90% was reached which was comparable to the accelerometer-based recognition performance also evaluated in the paper. Note that this study was performed using cross-validation.

In Ref. [83], we showed in a very limited case study that using RSS measured on a phone from nearby access points can be used to discriminate the three contexts: empty room, holding the phone and walking in proximity of the phone. Achieved accuracy for these experiments was approx. 70%.

Sigg et al. [81] presented a study showing the feasibility of phone-based gesture recognition using mobile phone WiFi RSS. In their study they leveraged RSS from all close by access points to determine a gesture performed by a user directly above the phone. The system discriminates 11 gestures or 4 gestures with accuracies of 50% and 72%, respectively. They evaluated a large number of cases and also conducted a feature subset selection identifying the mean, variance and range and the signal slope as most helpful features.

Finally, Abdelnasser et al. [2] presented a WiFi-based gesture recognition system. WiGest recognizes five types of in-air hand motion around a users' mobile device. The authors evaluated the system employing test beds with one and three access points and achieved 87% and 96% recognition accuracy, respectively. The system does not require a specific calibration phase. Instead it employs a unique preamble, i.e. a specific gesture, to tune the system for recognition. Similar to the work of Sigg et al. [81], to achieve a high accuracy the system requires the hand motion to be carried out in close proximity to the device.

## 8.3 ANALYSIS AND METHODS

This section begins with an exploration of the theoretical effect of activities on radio signals leading to an activity categorization. Next, we describe how the RFHAR system should be designed based on the best practices of classical Activity Recognition. Using the design concept and the activity categories we finally define an approach for RFHAR optimization and evaluation.

### 8.3.1 *Effect of Human Physical Activities on Radio Signal Strength*

In this section we derive a concept of the effect of human physical activities based on basic propagation theory. As explained in Chapter 2, the power level reflected in the received signal strength corresponds to the superimposition of all electromagnetic waves arriving at the same time at the receiver antenna (cf. Eq. 2). Before arriving at the receiver antenna the waves have interacted to different degrees with the surroundings which is reflected in each signal. Visualizing this concept, we are aware that a large area around two transceivers is filled with electromagnetic waves on many different propagation paths.

Introducing a subject, e.g. in line of sight between sender and receiver, will affect a number of these paths. Performing a periodic physical activity in this area leads to continuous changes of the affected paths. Using this understanding we discover that if a motion is repeated in the same way at the same location while nothing else in the environment changes, then it must affect the same paths. In fact, it should also lead to recognizable patterns in the RSS. Thus, depending on the character

of the physical activity the obtained RSS will look different. From this idea we can also grasp that the location at which a subject performs the activity must also have an impact. For instance, if the subject closes in on the receiver/transmitter the number of affected paths will increase.

Using this understanding we propose a coarse categorization of ADL with respect to device-free recognition as follows:

- **Category 1: Dynamic with changing location**
  Examples include activities such as “Walking” or “Crawling”, i.e. the movement of a large shape/volume in the area in which the electromagnetic field is extending. Depending on the location of the subject relative to the nodes these activities are likely to induce a strong signal fluctuation as the subject traverses the area affecting many propagation paths of a link. If the Radio Sensor is composed of multiple links, it is likely that multiple of these links are affected.

- **Category 2: Dynamic with static location**
  Examples include “Eating”, “Sitting and Typing”, food preparation and fitness exercises. Depending on the location of the transceivers these activities are likely to induce medium signal fluctuations as the conducted motion affects the same areas repeatedly. Thus, if the activities are periodic this must also be observable in the sensor signal. Due to the nature of the activity (locally restricted) it is likely that in a multi link Radio Sensor only a proportion of links will be affected.

- **Category 3: Static with static location**
  Examples include “Sleeping”, “Standing” and “Sitting”. These activities usually involve only little physical motion (e.g. breathing, swaying slightly while standing "still"). However, depending on the Radio Sensor configuration and the subject position they might still induce signal fluctuations. As the motion is very restricted (size and distance of moved object is rather small) this type of activity will generally yield the smallest effect in the radio signal.

### 8.3.2 *A priori Investigation*

To test our assumption of the different activity categories we conducted a number of activities between a single 3 m link of the Radio Sensor. Of each assumed activity category we selected two activities which are performed for 10 s each. Selected activities are “Walking” and “Crawling” for category 1, “Push-ups” and “Standing and Waving” for category 2. For category 3 “Lying” and “Standing” were selected. Activities with static location were performed in the middle of the link facing one of the nodes. Activities with changing location were conducted along the line of sight of the links. Figure 45 shows the raw data RSS for the recordings. An empty recording is provided as reference.

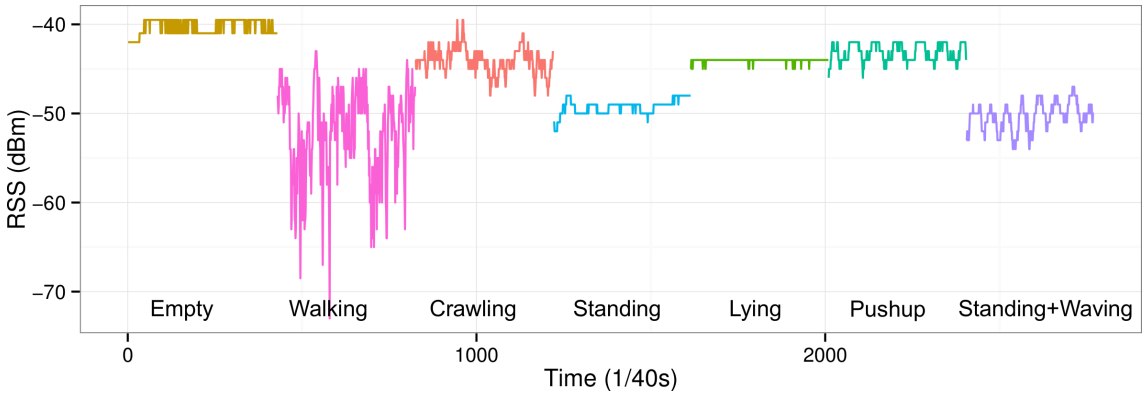



Figure 45.: RSS for six distinct activities, two of each assumed category. "Empty" denotes a reference measurement where no subject was in LoS.

| Empty | Walking | Crawling | Standing | Lying | Push-ups | Waving |
|---|---|---|---|---|---|---|
| -40.27 dBm | -52.18 dBm | -44.07 dBm | -49.27 dBm | -44.04 dBm | -43.11 dBm | -50.30 dBm |
| ±0.87 dBm | ±5.04 dBm | ±1.45 dBm | ±0.84 dBm | ±0.20 dBm | ±1.00 dBm | ±1.54 dBm |

Table 10.: Means and standard deviations of the recorded activities.

Table 10 shows the means and standard deviations of the activities. It can be observed that the "Empty" measurement had the strongest received signal strength and only little fluctuation. Activities conducted on the floor had a mean signal strength about 5 dBm higher than the activities "Walking", "Standing" or "Waving". For these activities the subject stood directly in the LoS and average signal power was the lowest. RSS standard deviation was maximal for "Walking", followed by "Standing and Waving", "Crawling", "Push-ups". The static activities "Standing" and "Lying" had standard deviations similar to the "Empty" measurement. As the absolute signal level is not a robust feature for context discrimination [45] we will focus on standard deviation for investigating how the results match to our previous assumption.

As expected, activities of category 3 induced the least signal fluctuation. However, "Crawling", which is defined as category 1 activity, had a standard deviation smaller than the activity "Standing and Waving". The measurements for all other activities followed the previous assumption in respect to their category. The reason why "Standing and Waving" influenced the signal more strongly is that it was conducted in LoS between transceivers. Put differently, it affected a larger number of multipaths as the surface presented to the transceivers was larger than for "Crawling". Since the fluctuation induced by "Crawling" is still very strong (the RSS range was larger than that of "Standing and Waving") we still consider the categorization as valid. However, the impact of the transceiver position in relation to the subject is important. Further on, the result indicates that a measure of variation alone is probably not sufficient for discriminating activity categories. For the above set of activities a metric of periodicity such as the frequency or similar could help to distinguish category 1 and 2 activities from each other.

### 8.3.3 *Activity Selection*

The previous section investigated the radio signature of a number of activities of different categories. From this brief analysis we found that the discrimination of activities of different categories may be achieved with good accuracy. On the other hand, the separation of activities within classes (e.g. for activities with very reduced motion) seems much harder and may require special or additional means such as location information. For this evaluation we selected a single activity from each activity category. Additionally, we define that all activities have the same base body pose (e.g. upright) so that the effect on RSS is not additionally influenced by a changed body area.

However, the selected activities should be natural and deliver information which could be helpful in a Smart Home system. Hence, we selected the following activities:

- Walking (cat 1)
- Standing and Waving (cat 2)
- Standing (cat 3)

The activity "Standing and Waving" seemed especially interesting as it may also allow explicit triggering of Smart Home actions.

### 8.3.4 *RFHAR System Design Considerations*

Classic Activity Recognition research has successfully demonstrated the application of an adapted machine learning pipeline architecture [8]. Thus, RFHAR should be developed by using this standard architecture as a template. This means that the system has to employ some kind of feature computation and classification steps. The system further needs a calibration or training phase prior to the training. Further on, the specific characteristics of the Radio Sensor may be harnessed using additional pipeline modules. As in classical Activity Recognition each of the processing steps has certain degrees of freedom, e.g. choice of features. Thus, in order to determine the optimal RFHAR implementation an additional optimization step should be conducted before evaluating the system. In this step various module configurations can be evaluated for suitability.

### 8.3.5 *Evaluation Strategy*

The RFHAR system should be developed with realistic conditions in mind. Among others, realistic conditions mean that the system provides a stable recognition rate over a longer period of time. This forbids the employment of typical cross-validation approaches as these involve training and testing from the same dataset. A better approach is to evaluate the system using independent datasets recorded over the course of multiple days at different times per day. Other aspects of

robust recognition include resistance to typical changes in the environment such as the opening and closing of doors or changes in the Radio Sensor infrastructure. Requiring realistic conditions also emphasises system practicability. Practicability involves the effort a user has to make to run the system. Assuming a completed installation, this relates to the amount of training time needed to get the system to a functional state. Hence, this is another aspect which should be evaluated.

To ensure RFHAR is evaluated against such realistic conditions albeit providing good recognition rates, we conducted RFHAR implementation in two steps.

1. **Optimization**: Define optimal parameter settings for RFHAR modules
2. **Evaluation**: Evaluate RFHAR in regard to overall goals

During optimization we used a single measurement campaign with several independent training and test sets to tune the RFHAR modules for optimal performance.

In the evaluation we used a larger number of measurement campaigns to inspect the performance of the tuned RFHAR system in respect to practical considerations.

For all investigations we used the following strategy to determine the performance: each single recording of an experiment is used as training set once and tested against all other recordings of this experiment on the same day. The performance of RFHAR for this training dataset is then the average of the recognition accuracies for the recordings used as test sets. This approach resembles a complete inversion of the cross-validation approach (where a single dataset is tested and all others are used for training) and makes achieving good results much more challenging. In order to allow the use of the multi class accuracy as benchmark, classes in all datasets were balanced before the evaluation.

### 8.3.6 *Summary*

In this section we have described how we approached the development of RFHAR in respect to the defined goals. The system itself is based on the machine learning pipeline as facilitated by Bao et al. [8] and suggested in Chapter 5. After the system was designed, it was optimized and evaluated in two separate steps. For optimization and evaluation a number of measurement campaigns were needed to provide enough data. The activities in these measurement campaigns were "Standing", "Standing and Waving" and "Walking". The data further respected specific requirements corresponding to the goals. For instance, datasets were recorded over multiple days with and without changed environments. For both, optimization and evaluation, the system was evaluated using a challenging test procedure in which only a single independent recording was used for training and all other recordings of this day were used for testing. In evaluation, considerations regarding the practicability such as training data length were further investigated. As metric for evaluation we selected multi class accuracy.

## 8.4 RFHAR SYSTEM

In this section we describe the design of the RFHAR device-free RSS-based Activity Recognition system. The RFHAR system uses a processing chain as introduced by Bao et al. [8] and discussed in Chapter 5. This chain is extended by modules (*Preprocessing*, *Fusing*) specifically introduced for the Radio Sensor. Nevertheless, the general work principle is unchanged: annotated groundtruth data must first be recorded and supplied to the system for training (offline). Thereafter, new data can be passed through the processing chain and is classified as a specific activity (online). We start by describing the overall system architecture. Then we present each system module in detail and describe the different parameters to implement this particular module. We will select the actual configuration in the next section.

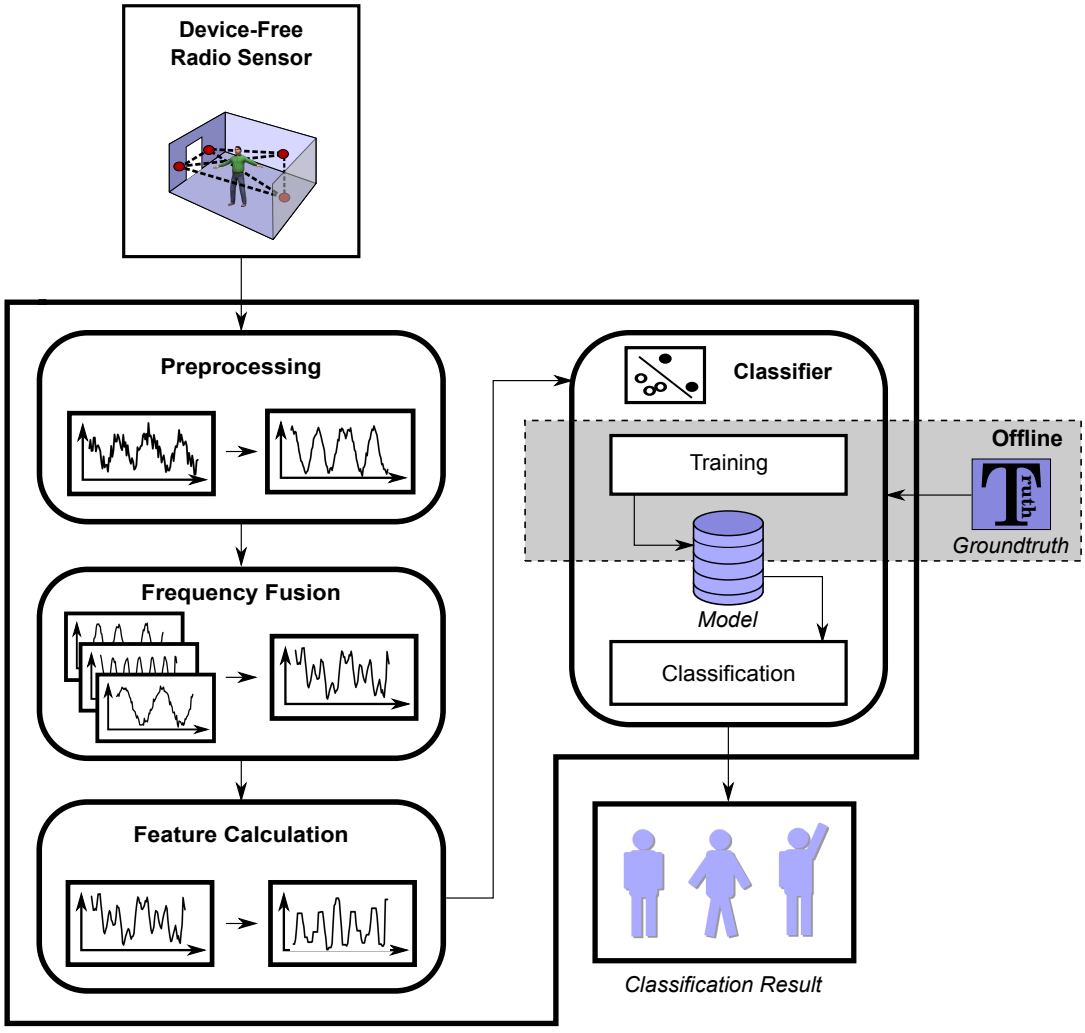



Figure 46.: RFHAR System Architecture

### 8.4.1 *Architecture Overview*

Figure 46 depicts the system architecture. The system has two general modes: the online phase and the offline phase. In the offline phase the *Training* function of the *Classifier* and the *Groundtruth* data is used. In the online phase only the *Classification* function from the *Classifier* is used. Apart from this difference the remaining system modules are used in both phases in the same way. The system modules are configured in a pipeline architecture. Each module receives input data from the previous module, transforms it and passes it on to the next module.

In RFHAR RSS values are continuously sampled by the *Radio Sensor*. This data stream is passed on to the *Preprocessing* module. *Preprocessing* cleans the data for further processing, for instance it may smooth the data using moving average filtering. Thereafter the *Frequency Fusion* module combines measurements of the same wireless link on multiple frequencies. This combined data stream is then handed over to the *Feature Calculation* which applies some mathematical function over a number of preprocessed measurement values to increase the separability of the conducted activities. This feature data is then passed on to the *Classifier*. If in offline mode the classification algorithm uses the provided *Groundtruth* to build an internal model of the relation between measured data and classes. If in online mode, the classification algorithm uses this internal model to classify the measured data. The result of this step is the *Classification Result*.

### 8.4.2 *Preprocessing*

The received signal strength acquired by COTS transceivers is a noisy, low resolution signal power information. In this context, noise means signal changes created by other parameters influencing the RSS (cf. Chapter 4) and not through the subject's activity. As an example, consider the RSS measurements for "Empty" in Figure 45. No activity took place in the room, hence we assume a constant signal level. But instead, the measured signal fluctuated frequently on the order of +/-1dBm around the mean. Hence, we can characterize the noise in RSS as high frequency and low magnitude compared to the signal fluctuations introduced by human activities. In signal processing, filtering is employed to reject or attenuate specific frequencies and extract only the signal of relevance [32]. In this module, we consider two such filters: the moving average filter, a classical low pass filter and the amplitude filter, a filter specifically designed for the observed noise characteristic. To compare the effect of the filters, we also include the option "None" which forwards the data to the next module without modification. The following list summarizes the *Preprocessing* options:

- **None**: Data is passed on as received by the *Radio Sensor*.

- **AmplFilter**: Amplitude filter which removes 1dBm fluctuations around the mean value as observed in Figure 45. Note that the Amplitude filter uses the same window size and shift width as the *Feature Calculation* module.

$$\text{AmplFilter}(t) = \begin{cases} x(t): & \text{if abs}(x(t) - \mu(X)) > 1 \\ \mu(X): & \text{else} \end{cases} \tag{32}$$

  Here, $\mu(X)$ is the mean over the current sample window of $x(t)$.

- **MvgAvg**: Moving average filter which smoothes data based on the mean over three consecutive measurements.

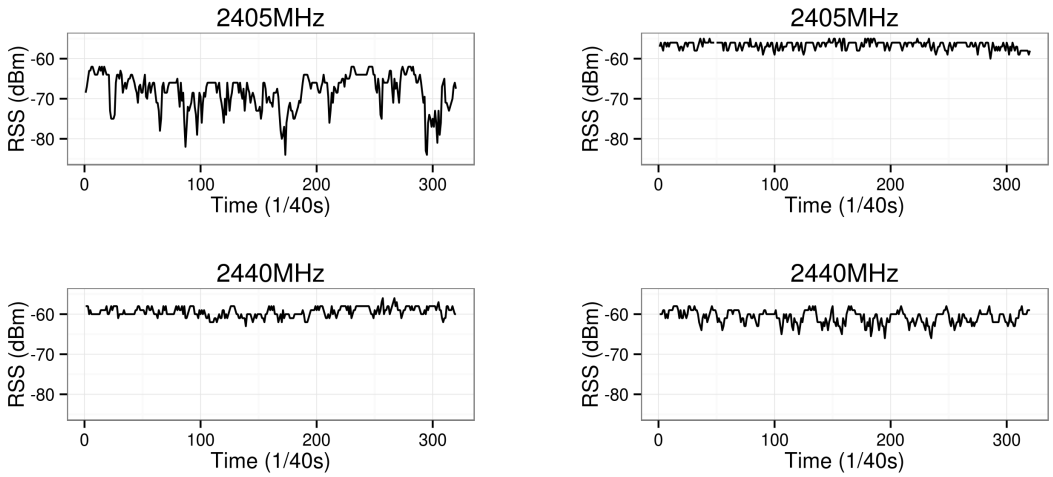
$$\text{MvgAvg}(t) = \frac{1}{3}\sum_{i=0}^{2} x(t-i) \tag{33}$$

### 8.4.3 *Frequency Fusion*

This modules combines various frequency measurements of the same link to a single measurement value. The received signal strength is a magnitude of very limited resolution. For the developed testbed, which uses the NXP Jennic 5139 SMA module, measurements ranged from -98 dBm to -11 dBm with a resolution of 2 dBm-3 dBm (cf. Appendix A). This resulted in only 46 different signal levels. In practice only a fraction of these signal levels are observed. Additionally, using the *Preprocessing* module may reduce this resolution even further. To improve sensor sensitivity the use of multiple frequencies can be helpful. In Chapter 4, we have shown that the use of multiple frequencies improves the recognition accuracy as different frequencies, even for the same link, respond slightly differently depending on their wavelength. Thus, the intuition behind this module is to leverage this observation: by combining the information of multiple frequencies into a single signal we remove redundancy (multiple signals which measure the same physical space) and instead explicitly add this redundancy as knowledge to the data (single signal with higher resolution which measures the same physical space).

Further on, the combination of frequencies into a single signal may help to make RFHAR more robust. Consider Figure 47 as an example: the figure shows two recordings of the activity "Standing and Waving" performed on the same day, by the same subject at the same location. If a classifier is trained using separate frequency data with the data from 10:30am (left) and tested on the data from 5:30pm it may have difficulties identifying the activity. If the signals were combined (e.g. summed), the two recordings would still not be identical but the difference would be reduced.

No best practice for RSS frequency combination is available. Hence, we consider a naive and a more sophisticated approach. To allow comparison of the effect of combination we also include the “None” configuration which does not combine the measurements. In the following equations, $x_i$ denotes an RSS measurement on frequency $i$ at a specific time. $F$ denotes the number of frequencies measured and




(a) Aug 28, 2014, 10:30am (b) Aug 28, 2014, 05:30pm

Figure 47.: Two measurements for the activity "Standing and Waving" at two different times on the same day.

$\mu(X_i)$ denotes the mean over a number of measurements for a specific frequency $i$. Note that only frequencies belonging to the same wireless link are combined. The following list summarizes the *Frequency Fusion* options:

- **None**: Measurements on different frequencies are not combined.
- **Mean**: Averaging of frequencies is the most straight forward approach. However, this approach does not consider that an activity may have the opposite effect on two frequencies. I.e. when one frequency is attenuated, the other may be amplified possibly cancelling out each other. Thus, chances that this approach improves recognition performance are considered low.

$$\text{mean}(x) = \frac{1}{F}\sum_{i=0}^{F-1} x_i \tag{34}$$

- **Mean Corrected Absolute Sum (MCAS)**: By removing the mean $\mu(X_i)$ from $x_i$ and computing the absolute value prior to summation, we ensure that fluctuations add up accordingly across frequencies. Note that the computation of $\mu(X_i)$ uses the same window size and shift as the *Feature Calculation*.

$$\text{mcas}(x) = \sum_{i=0}^{F-1} \text{abs}(x_i - \mu(X_i) \tag{35}$$

### 8.4.4 *Feature Calculation*

This module derives features based on measured raw data. Thus, this module has three configuration options. First, the number of values (feature window size) used

for the feature computation. Second, the mathematical function computing the feature. And third, as Bao et al. [8] showed that overlapping feature windows are helpful for recognition, the window shift width. Based on our previous work [73] in which we achieved decent cross-validation results using 1 s windows, we chose the window lengths and shift widths shown in Table 11. If window size and shift width are identical, windows are not overlapping.

| Window Size (s) | Shift Width (s) |
|---|---|
| 0.5 | 0.25 |
| 0.5 | 0.5 |
| 1.0 | 0.5 |
| 1.0 | 1.0 |

Table 11.: Window sizes and shift widths

In the following, we select a number of features to investigate in the optimization. Note that in the equations $X$ denotes the vector which holds all values in the current window and $|X|$ is the number of elements in this vector.

Although it is likely to be strongly affected by other environmental changes [45] we include the mean for comparison.

$$\text{mean}(X) = \frac{1}{|X|} \sum_{i=0}^{|X|-1} x_i \tag{36}$$

Statistical measures of dispersion are most promising for good recognition performance. While such measures are relatively robust against environmental dispersion over time [45], they are also well suited to capture activity characteristics. A prime example of such a measure is the standard deviation (Std). In addition, we include the range and the inter quartile range (IQR). In contrast to standard deviation, the range is especially prone to outliers which could be helpful in respect to the low resolution of the sensor measurements. On the other hand, the IQR is especially resistant against outliers and may act as additional noise filter [23]. The IQR is the difference between the third and first quartile ($Q_x$) of a sample.

$$\text{Std}(X) = \sqrt{\frac{1}{|X|-1} \sum_{i=0}^{|X|-1} (x_i - \text{mean}(X))^2} \tag{37}$$

$$\text{Range}(X) = \max(X) - \min(X) \tag{38}$$

$$\text{IQR}(X) = Q_3(X) - Q_1(X) \tag{39}$$

A fourth measure of dispersion is the mean absolute deviation (MAD). Similar to the IQR the MAD is relatively robust against outliers.

$$\text{MAD}(X) = \frac{1}{|X|} \sum_{i=0}^{|X|-1} \text{abs}(x_i - \text{mean}(X)) \tag{40}$$

Frequency related features could also be helpful to remove noise and to extract the frequency of actual activities. For this reason we also include the FFT as a feature. In order to keep the feature space dimension low we use the sum of the FFT after detrending using the mean. I.e.

$$\text{Sum\_FFT}(X) = \sum_{f}^{\frac{|X|}{2}} \text{FFT}_f(X - \text{mean}(X)) \tag{41}$$

Due to the limited resolution of the measurements, specifically developed features could help to increase system performance. Comparing the three activity categories and the selected example activities in Fig. 45, we find two main differences between activities: 1) magnitude of fluctuation and 2) frequency of fluctuation. Magnitude of fluctuation means how strong the signal changes within a specific time period. Frequency of fluctuation describes how often the signal changes within a specific time period. The frequency may be determined using the FFT. An alternative approach is to consider the speed of signal change. We define the speed as the number of intermediate signal steps between a $\text{RSS}(t_0)$ measured at some initial time stamp $t_0$ and an $\text{RSS}(t_0 + dt)$ measured some time later. To capture this information we develop a new feature: LUnique. LUnique is the number of unique elements in the current measurement vector $X$. Figure 48 shows the pseudocode for the computation of the LUnique feature.

**Data**: X
**Result**: Number of unique elements in X
$Y \longleftarrow \emptyset$;
**foreach** *i in X* **do**
  **if** *i not in Y* **then**
    put *i* in Y
  **end**
**end**
**return** ***length***(*Y*)

Figure 48.: Computation of LUnique

### 8.4.5 *Classifier*

As demonstrated by Bao in 2004 [8], machine learning algorithms are well suited for the recognition of human activities using accelerometers. In 2010, Gordon et al. [27] have studied the aspects of introducing novel sensors into Activity Recognition. One of their advises was to consider multiple recognition algorithms. We have followed up on this advise in Scholz et al. [73]. Therein we evaluated three different classifiers on a device-free Activity Recognition dataset: k-Nearest Neighbors (k-NN) [18], Naive Bayes [48] and Decision Tree C4.5 [62]. The results were very promising but inconclusive showing very different behavior among the

classifiers. In our cross-validation evaluation, the k-NN (k=1) achieved the best recognition performance but in a training/test split evaluation Naive Bayes and the C4.5 were superior.

Thus, for RFHAR we select all of these classifiers as possible candidates. In addition, we include the Support Vector Machine (SVM) algorithm [12]. The SVM algorithm has a number of features which make it interesting to test for device-free Activity Recognition. Most importantly SVMs are said to be robust even to non-optimal features as they maximize margins between classes using their kernel functions resulting in a very good generalization ability [3]. Shi et al. [78] also reported from improved results when employing a SVM for SDR-based Activity Recognition. Thus, we will also include the SVM classifier as candidate for the *Classification* module in the next section.

### 8.4.6 *Summary*

Table 12 summarizes the configuration options for each RFHAR module.

| Module | Configuration |
|---|---|
| *Preprocessing* | None, MvgAvg, AmplFilter |
| *Frequency Fusion* | None, Mean, MCAS |
| *Feature Calculation*: Feature | Mean, Sum_FFT, Range, IQR, MAD, LUnique, Std |
| *Feature Calculation*: Window Size | 20 (0.5 s) or 40 samples (1 s) |
| *Feature Calculation*: Window Shift | 0% or 50% |
| *Classifier* | C4.5, Naive Bayes, k-NN, SVM |

Table 12.: Summary of RFHAR modules and configuration options

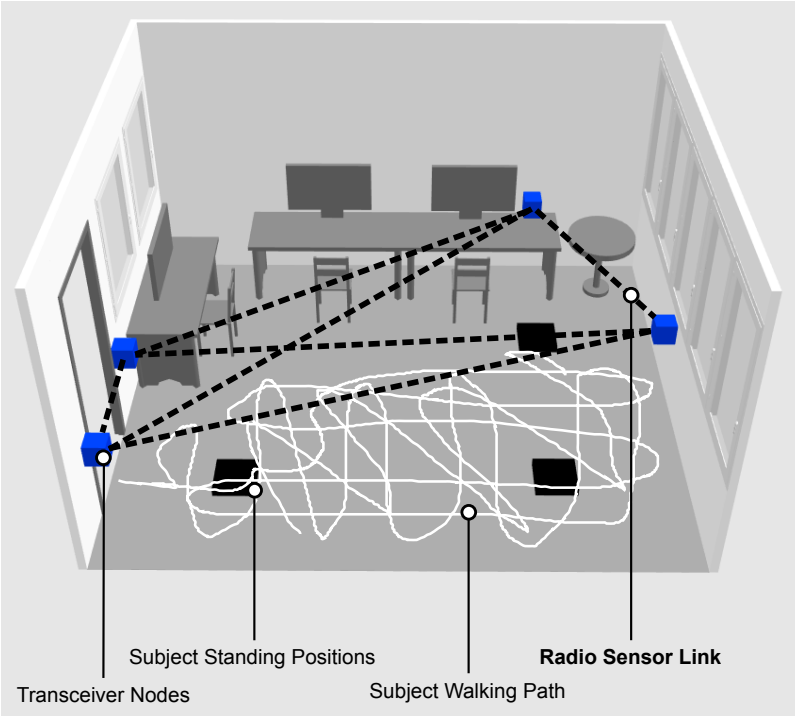
## 8.5 MEASUREMENT CAMPAIGNS

To enable RFHAR optimization and evaluation, four measurement campaigns were conducted. The first campaign (R1S1) took place in a different room (Office1) than the other three campaigns which were conducted in the Office2 room. In all campaigns the Radio Sensor implementation described in Appendix A was employed. Thus, each room was outfitted with four Jennic 5139 IEEE 802.15.4 transceivers. These measure received signal strengths with 40 Hz on three carrier frequencies ($f_c$ = 2405 MHz, 2440 MHz, 2480 MHz) between all nodes. In all campaigns, two of the transceivers were placed left and right of the entrance to the room while the other transceivers were placed in the far left corner away from the room entrance. In the following subsections we briefly describe each campaign.

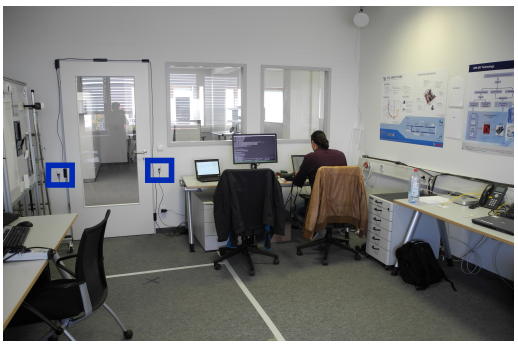
### 8.5.1 *R1S1*

This measurement campaign includes only the activities "Walking" and "Standing". Nevertheless, it was included for the evaluation since it took place in a

different room compared to the other campaigns. The testbed and activity sequence is identical to the LoS-Cross evaluation in Chapter 6. For the readers convenience we repeat the most relevant part of the description here.




(a) Schematic of installation and experiment



(b) Image of transceivers installed left and right of the door

(c) Image of transceivers on the window front and left wall of the room

Figure 49.: Illustration of the R1S1 testbed.

The testbed was installed in a 5.50x4.60 m office room with gypsum walls. Nodes were placed at 0.80 m height either on stands (node in front of window) or directly on the wall (above the desk; right and left of door) we refer to this topology as Topology ID:1. Recordings were performed with open and closed door. The subject (Subject ID:1) conducting the experiments was 1.76 m tall and weighed 72 kg. An illustration of the R1S1 testbed is given in Figure 49. The broken black lines indicate the LoS of the wireless links between the transceivers which are represented by cubes. The conducted activity sequence was a mix of "Standing" and "Walking" activities. The white line in the picture illustrates the walking

path which covers the complete room. The "Standing" activity was conducted at three distinct positions (black squares on the room floor in the figure). The activity sequence was as follows:

1. Walking in the room and stopping at the top right position in Fig. 49
2. Standing at this position for 10 s
3. Walking in the room for a while and stopping at the lower right position
4. Standing at this position for 10 s
5. Walking in the room for a while and stopping at the position closest to the door
6. Standing at this position for 10 s

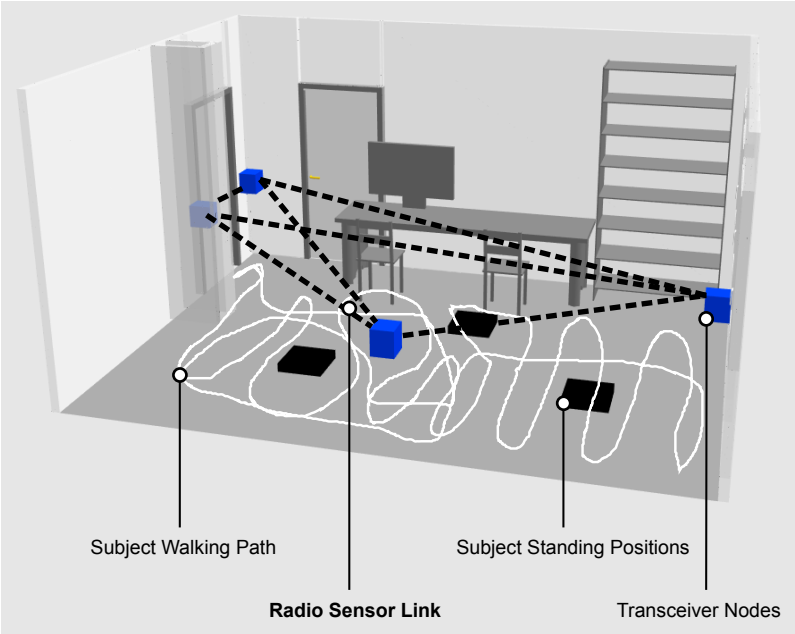
In the first iteration the room door was open. Afterwards, the sequence was repeated with a closed door. The average length of a single evaluation dataset was 2.5 min. A total of 28 recordings were made on four different workdays within a single week. Recordings were conducted between 10am and 6pm. The total recording time was 70 min.

### 8.5.2 *R2S2*

In this measurement campaign the activities "Walking", "Standing" and "Standing and Waving" were recorded in the Office2 testbed. Office2 was a 2 person office room at the university. It had an area of 4.04x5.33 m size, gypsum walls on three sides and a window front with a concrete wall. It had a bare concrete ceiling at a height of 2.70 m. The room featured typical pressboard furniture, two office chairs and two computers. Figure 50 shows a schematic illustration of the room. The figure shows the walking path of the subject (white line), the LoS of the wireless links (black broken line), the transceivers (four cubes) and the three positions at which the activity "Standing" and the activity "Standing and Waving" was performed (black squares on the floor). As in the R1S1 installation the transceivers were installed at 0.80 m height. They were located right and left of the main door, on the window front and on the wall right of the entrance. Transceivers were located on columns (window front, wall) or glued to the wall (left and right of door). We refer to this transceiver installation as Topology ID:2. The experiments were conducted by a male subject which was 1.70 m tall and weighed 68 kg (Subject ID:2).

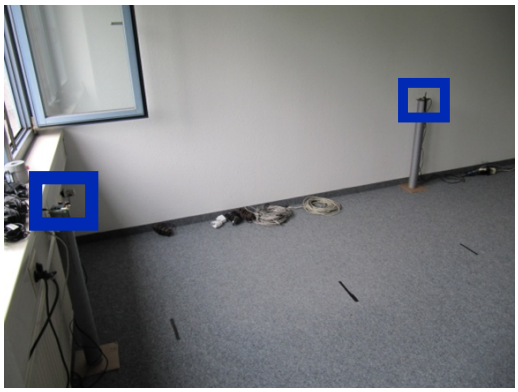
The recordings were performed on a total of three days. Of which the second day was a Saturday. Two recordings were performed per day with a break of 30 to 60 minutes between recordings. The recorded activity sequence for each of the six datasets was as follows:

1. Walking (20 s)
2. Standing at the position closest to the entrance (20 s)
3. Standing and Waving at this position (20 s)
4. Walking (20 s)
5. Standing at the position closest to the window front (20 s)
6. Standing and Waving at this position (20 s)




(a) Schematic of installation and experiment

(b) Image with transceivers left and right of the door highlighted



(c) Image of transceivers on the wall and the window front highlighted

Figure 50.: Illustration of the Office2 Testbed for the R2S2 campaign.

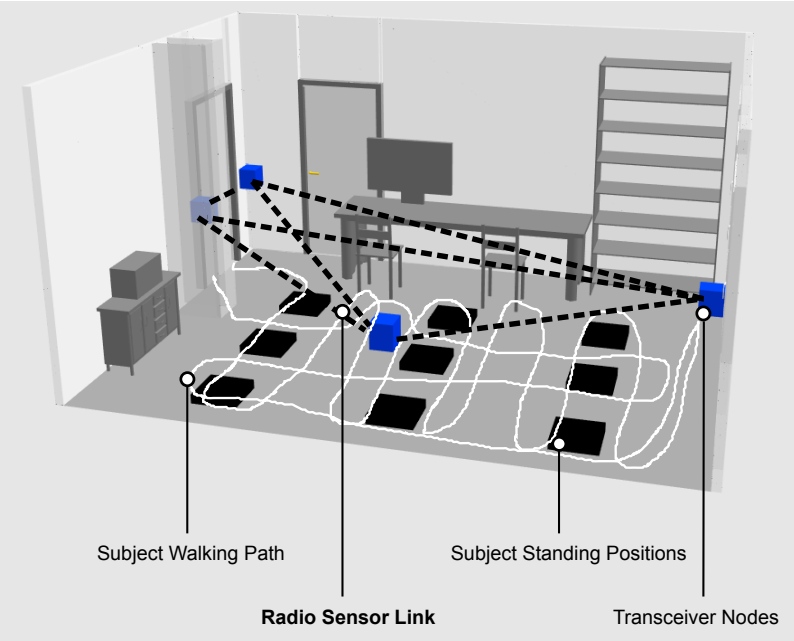
7. Walking (20 s)
8. Standing at the position closest to the table (20 s)
9. Standing and Waving at this position (20 s)

Each dataset was 3 min long. The combined time of the recordings was 18 min.

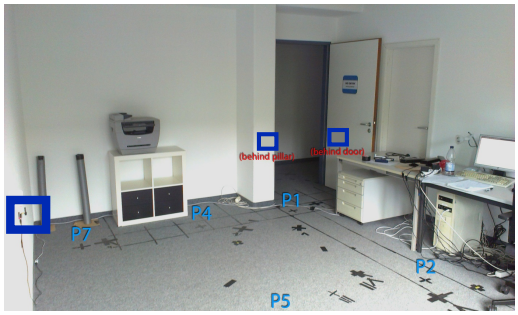
### 8.5.3 *R2S3A*

In this measurement campaign datasets were recorded in Office2. However, transceivers were all attached to the walls (Topology ID:3). Also the activities of category 2 and 3 were performed at nine distinct locations and in two orientations. The visualization of the experiment in Figure 51 shows the positions of these

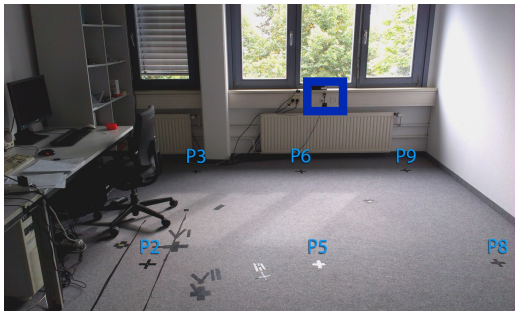
locations. A male subject which was 1.66 m tall and weighed 64 kg (Subject ID:3) performed the activity experiments.




(a) Schematic of installation and experiment




(b) Image with positions enumerated and transceivers around door and on the wall highlighted




(c) Image with positions enumerated and transceiver at the window front highlighted

Figure 51.: Illustration of the Office2 Testbed for the campaigns R2S3A and R2S3B.

All datasets in this campaign were recorded on a single week day. Recordings were taken between 10am and 5.30pm. Approx. two recordings were taken per hour. Each activity of category 2 or 3 was conducted in two orientations: 1) facing the table and 2) 90° rotated facing the window front. The nine activity locations are depicted in the figures 51b and 51c by PX's. The recorded activity sequence for each of the 9 datasets was as follows:

1. Walking in the whole room (180 s)
2. For each position PX=$P1, \ldots, P9$:
   a) Standing at PX facing table (9 s)
   b) Standing at PX facing window (9 s)
   c) Standing and Waving at PX facing table (9 s)

d) Standing and Waving at PX facing window (9 s)

The length of a single experiment dataset was 8.4 min. The combined time of all recordings in the campaign was 75.6min.

### 8.5.4 *R2S3B*

The R2S3B campaign was a larger version of R2S3A. It spanned six days in total. On each day six recordings were conducted. The conducted activity sequence and execution times are identical to the recordings in R2S3A. However, one important difference is that the Radio Sensor topology was changed slightly (Topology ID:4). The transceiver attached to the wall was moved 30 cm away from the wall and into the direction of the door and placed on a styrofoam column. Figure 52 shows the changed position of the transceiver.

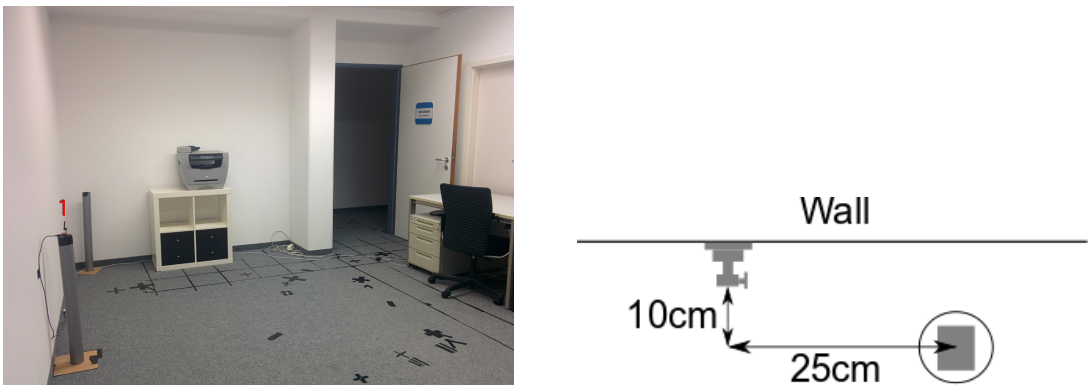



Figure 52.: Topology change: relocation of transceiver from wall to column

The combined time of all datasets recorded on a single day was 50.4 min. The combined time including all six days was 302.4 min. The total time span in which all six full day recordings were made was 12 days.

### 8.5.5 *Summary*

Table 13 shows a summary of the different measurement campaigns that were recorded over a period of more than one year. Three of four measurement campaigns were conducted over multiple days allowing evaluation of training data age. The datasets from all measurement campaigns except R1S1 include activities of all three categories. Combining the data from all campaigns we have over 7.5 h of activity data including two rooms of different size and four different Radio Sensor topologies (Topo.ID). Using the R1S1 data we can further investigate the effect of different door states more closely.

| Campaign | Timespan | Days | #Rec | Subj.ID | Acts | Length(min) | Rooms | Topo.ID | Door |
|---|---|---|---|---|---|---|---|---|---|
| R1S1 | 25.10.-30.10.2013 | 4 | 28 | 1 | 2 | 70 | Office1 | 1 | open&closed |
| R2S2 | 17.10.-21.10.2013 | 3 | 6 | 2 | 3 | 18 | Office2 | 2 | open |
| R2S3A | 29.08.2014 | 1 | 9 | 3 | 3 | 75 | Office2 | 3 | open |
| R2S3B | 28.10.-04.11.2014 | 6 | 36 | 3 | 3 | 302 | Office2 | 4 | open |
| Total | >1 year | 14 | 79 | 3 | 2-3 | 465 | 2 | 4 | open/closed |

Table 13.: Summary over measurement campaigns

## 8.6 OPTIMIZATION

In this section, we describe the optimization of the RFHAR system. Therefore, we evaluated the parameters in each of the modules for their optimal performance against the R2S2 measurement campaign using the procedure described in the analysis section. We selected the optimal parameters using the median accuracy and inter quartile range of the accuracy as primary and secondary criteria, respectively. At the end of the section we summarize the optimal RFHAR settings, which we used in the next step for a thorough RFHAR evaluation. We tested each system parameter by varying its settings and fixing the other parameters to default values (cf. Tab. 14).

| RFHAR Module | Default Parameter Setting |
|---|---|
| Preprocessing | Moving Average |
| Fusion | MCAS |
| Feature Calculation: Window Size | 40 |
| Feature Calculation: Window Shift | 40 |
| Feature Calculation: Feature | LUnique |
| Classifier | Naive Bayes |

Table 14.: Default configuration for the RFHAR optimization

### 8.6.1 *Preprocessing*

The *Preprocessing* module has three possible configurations: “None”, amplitude filter (AmplFilter) and moving average filter (MvgAvg). We evaluated the performance of each using the R2S2 measurement campaign. The box plot in Figure 53 shows the results of the evaluation. While the amplitude filter obviously removed too much information by dropping signal fluctuations of 1dBm around the mean, the moving average filter and the raw data delivered similar results. The moving average filter outperformed the “None” configuration by 2% median accuracy. The moving average filter configuration was also superior regarding the inter quartile range.

### 8.6.2 *Frequency Fusion*

The *Frequency Fusion* module has three possible configuration options: “None”, mean corrected absolute sum (MCAS) and the average across all frequencies

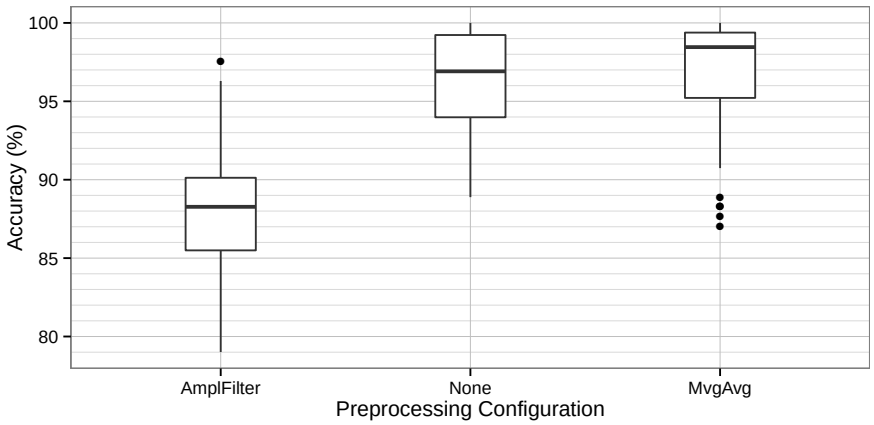



Figure 53.: Influence of the *Preprocessing* configuration on the accuracy.

(MEAN). Figure 54 shows the performance of each configuration using default parameters. While MCAS had a larger spread among the achieved recognition accuracies, the median accuracy was 2% points better than for the MEAN fusing. Both MCAS and MEAN outperformed the configuration without fusing.

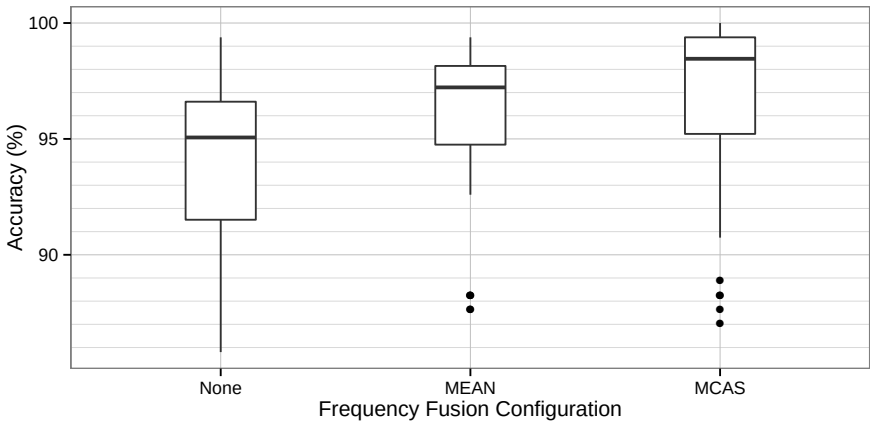



Figure 54.: Influence of the *Frequency Fusion* configuration on the accuracy.

### 8.6.3 *Feature Calculation*

The feature calculation has two different configuration options: the feature window and the actual feature.

#### *Feature Window Configuration*

A feature window contains the raw data instances passed on to the feature calculation. It may be configured regarding its length and the degree of overlap. Figure 55 shows the RFHAR performance when utilizing different feature windows. All configurations achieved a median accuracy of 95% or higher. The best median

performance was achieved using the maximal window size (40) without overlap. On the other hand, the accuracies using a smaller window size (20) showed less fluctuations and may be considered more stable. The good performance of the 40/40 configuration may have the followings reasons: 1) The frequency of the "Standing and Waving" activity conducted by the subjects was typically close to 1 Hz. Thus, 40 samples captured exactly one iteration of this activity. This would also explain the larger inter quartile range compared to the 20 sample windows: If the 40 sample window was not well aligned with the activity the performance decreased, leading to a larger IQR. In contrast, the short windows always captured only a part of the activity leading to a reduced but more stable recognition rate. 2) Activities started directly after begin of the recording. As recordings were always an even number of seconds long an overlapping window probably provided no advantage. However, in a freely running system overlapping windows may provide better results as activities are not so perfectly aligned. Nevertheless, achieving a higher accuracy is more likely using the maximal window configuration therefore we choose this as optimal parameter setting.

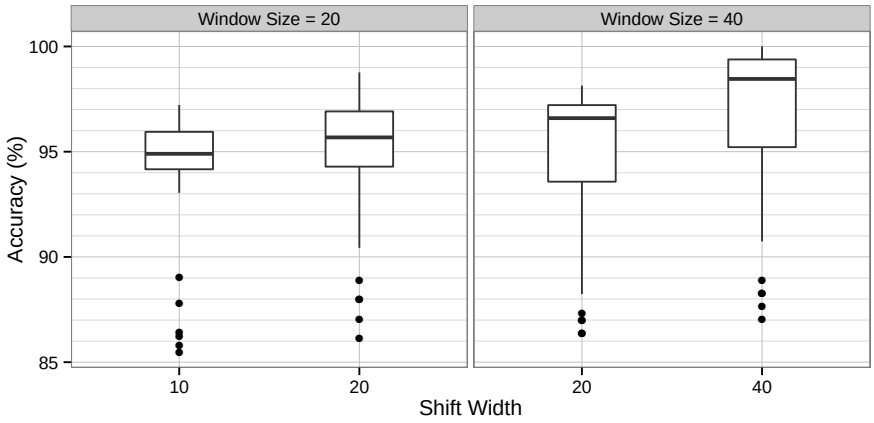



Figure 55.: *Feature Calculation*: Influence of window size and shift width on accuracy.

*Feature*

In the RFHAR design we have defined a number of possible features which could be beneficial for discriminating activities. Some of the features have been chosen as they have been frequently used by the research community e.g. for localization (standard deviation) while others could exploit specific characteristics of activities such as their periodicity (e.g. FFT). Figure 56 shows the result of the feature evaluation. Based on the performance, the features may be categorized in two groups: 1) Features which achieved approximately 95% median accuracy. 2) Features which achieved a median accuracy above 95%. The first group contains IQR, Std, MAD, Mean and Range. The Mean feature performed better than expected as it is known for its susceptibility to environmental changes. It seems

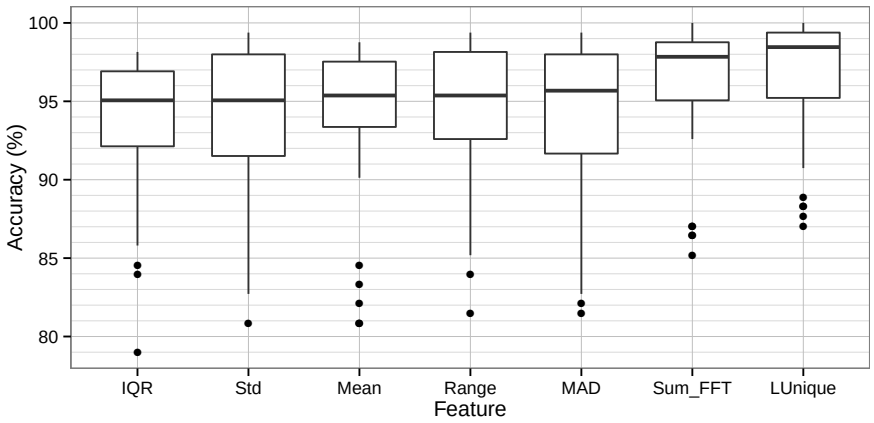



Figure 56.: *Feature Calculation*: Feature type vs accuracy.

that although the R2S2 campaign is spread across three days the environment was stable enough to allow good recognition accuracy using the Mean feature. On the other hand, the standard deviation (Std) which is frequently used in device-free Presence Detection [45] and for Activity Recognition [73] was expected to perform better. Obviously the periodicity of activities as reflected by the Sum_FFT was superior to the information captured by the first group of features. However, the LUnique feature outperforms the Sum_FFT because it covers both the periodicity and the actual fluctuation of the signal. Thus, for systems utilizing only a single feature and low resolution RSS measurements this seems to be an optimal fit.

### 8.6.4 *Classifier*

For RFHAR we selected four classifier options: Naive Bayes, Support Vector Machine (SVM), Decision Tree C4.5 and the k-Nearest Neighbors. Each classifier has parameters for tuning these for the given data (besides the actual training). However, especially the SVM and k-NN algorithms have specific parameters which are known to greatly affect the recognition performance [10]. Hence, we first estimate the optimal SVM kernel and the optimal k parameter for the k-NN (Tuning). Then we compare all four classifiers (Comparison).

#### *Tuning*

For tuning we utilized the aforementioned evaluation strategy (hold out validation using independent training and test sets) and test each parameter setting with the default parameters.

K NEAREST NEIGHBORS For the k-NN algorithm [18] the mandatory parameter is the number of $k$ neighbors to be considered for choosing a class. An approximation of a good $k$ is $k = \sqrt{N}$ where $N$ is the number of training samples [22]. Thus, in our case a good $k$ would be 22 as we have 466 training instances with the

default window size and shift. However, the optimal $k$ may not follow this rule and may alternatively be discovered by cross-validation or similar approaches [22]. Thus, we choose to evaluate $k = 1, .., 3 * \sqrt{N}$ to decide on the optimal $k$.

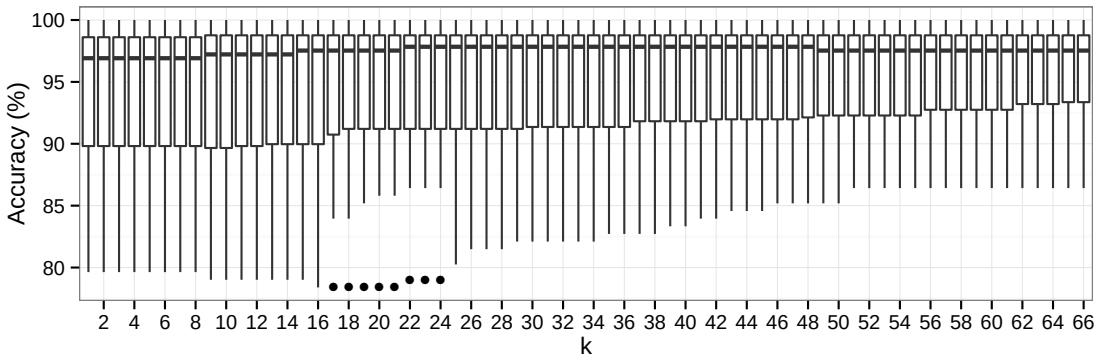



Figure 57.: Evaluation of RFHAR with k-NN for various k

Figure 57 shows the evaluation for $k = 1..66$. We can observe that $k = 22..48$ achieved a median accuracy of approx. 97%. After $k = 48$ the performance decreased. For choosing the optimal $k$ we enlist the exact accuracies in Table 15.

| k | Median Accuracy (%) | Standard Deviation Accuracy (%) |
|---|---|---|
| ⋮ | ⋮ | ⋮ |
| 21 | 97.53 | 5.74 |
| 22 | 97.84 | 5.63 |
| 23 | 97.84 | 5.63 |
| ⋮ | ⋮ | ⋮ |
| 47 | 97.84 | 4.86 |
| **48** | **97.84** | **4.84** |
| 49 | 97.53 | 4.83 |
| ⋮ | ⋮ | ⋮ |

Table 15.: Accuracies and standard deviation for different k.

We observe that while median accuracy is identical the standard deviation fluctuated, indicating that the accuracy for some of the tested instances improved with increasing $k$. Hence, we chose $k = 48$ as it showed the highest accuracy and the lowest dispersion.

SUPPORT VECTOR MACHINE The SVM [12] can be configured in respect to the non-linear function which transforms the training data to a higher dimension to make it linearly separable. This function, the SVM kernel, is chosen by the user while the function's parameters are usually optimized during the training phase. RFHAR utilizes the Orange data mining framework [21] for SVM training. Orange implements a grid search using 5-fold cross-validation to identify optimal settings for these parameters. Using the framework we tested the following four common kernel functions K [33] for their RFHAR performance:

- Linear: $\mathrm{K}(x_i, x_j) = x_i^T x_j$

- Polynomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$
- Radial Basis Function(RBF): $K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2), \gamma > 0$
- Sigmoid: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$

Here, $x$ is the training vector and $\gamma$, $d$ and $r$ are kernel parameters.

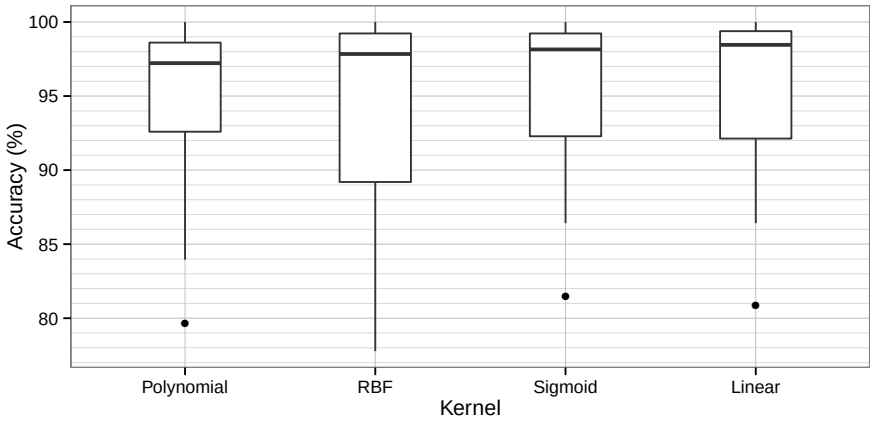



Figure 58.: Influence of the SVM kernel on accuracy.

Figure 58 illustrates the results of the performance evaluation. The performance of all kernel functions is relatively close and ranged from 97.2% to 98.4%. Interestingly, the RBF kernel which usually provides the top performance only ranked second and has the largest inter quartile range. Thus, we selected the Linear kernel as best configuration. It reached the best median accuracy (98.4%) and has a very limited problem-specific parameterization.

*Classifier Performance Comparison*

In this section we evaluate all four classifiers for their performance on the R2S2 campaign. Figure 59 shows the performance of each classifier. The decision tree algorithm (C4.5) had the lowest performance (89.5%) but the highest standard deviation. k-NN achieved a good median performance of 97.8%. SVM and Naive Bayes had the same accuracy (98.4%) but different standard deviations (Bayes: 4.3%, SVM: 5.3%). For this reason we chose Naive Bayes as the optimal classification algorithm for RFHAR.

### 8.6.5 *Optimal System Configuration*

The optimal RFHAR configuration employs the moving average filter and MCAS frequency fusion. It further uses the LUnique feature on a window size of 40 instances with non-overlapping windows and the Naive Bayes classification algorithm.

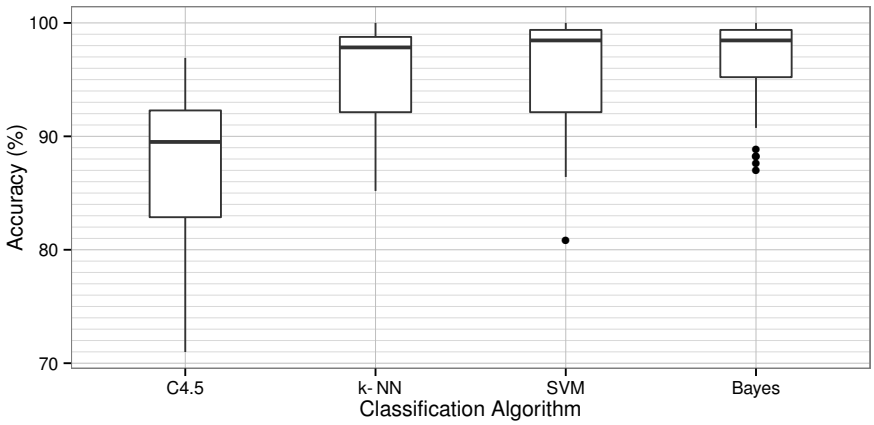



Figure 59.: Influence of the *Classifier type* on accuracy.

## 8.7 EVALUATION

In this section we evaluated RFHAR using data from the different measurement campaigns. As described in the introduction we are especially interested in the robustness of RFHAR regarding a number of aspects. To get a general understanding of the RFHAR performance the section is started with an overview of recognition performance for all provided measurement campaigns. For multi day campaigns we evaluate each single day. Thereafter, we look more closely at the aspects of training dataset age in respect to classification performance, the amount of training data and the effect of changing the Radio Sensor topology and the room geometry.

### 8.7.1 *System Performance across Measurement Campaigns*

Using the optimal RFHAR configuration we evaluated each day of each measurement campaign. This investigation explored RFHAR recognition performance when testing the system on the same day that it was trained on. For this evaluation we split the R1S1 campaign based on the door state. For evaluation the aforementioned evaluation strategy was employed. Thus, we used each single recording of each dataset as training dataset once and test against all other recordings of this day. This resembles a harsh test for the system as all test sets are independent and not generated from the same base dataset. Figure 60 shows the results of the evaluation. We observed that the achieved average accuracies range from 100% down to around 80%. The best accuracies (95-100%) were achieved for the R1S1 datasets. This could be expected as these include only two activities of the categories 1 ("Walking") and 3 ("Standing") which are most different in their characteristics. However, the second best results were achieved for the R2S2 dataset (>95%) which contains recordings for all activity categories. Lastly, R2S3A and R2S3B achieved accuracies ranging from 81-94%. Whereby R2S3B evaluations showed a higher spread in accuracy. Thus, the difference between R2S2 and R2S3A and B datasets

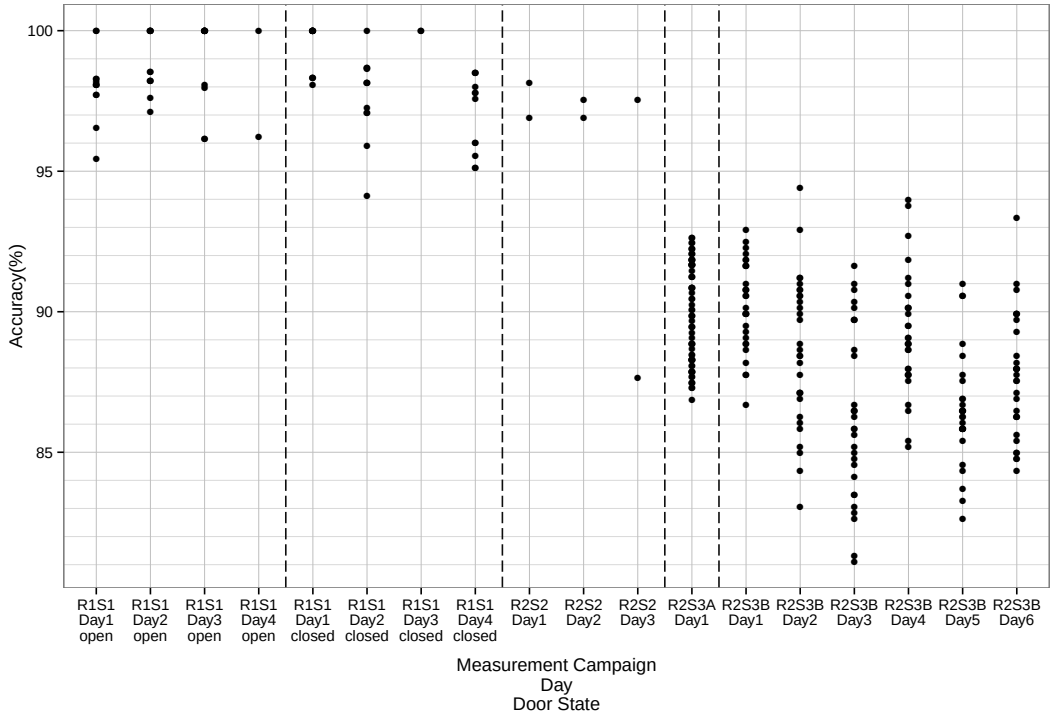



Figure 60.: RFHAR Performance across individual days of each measurement campaign. Each dot in the plot shows the result of single evaluation using a single recording for training and an independent recording on the same day as test. The broken line separates the different datasets.

were roughly 10%. Investigating this matter further we generated the confusion matrices for R2S2 and R2S3A (cf. Fig. 61).

| Truth \ RFHAR | Walking | Waving | Standing |
|---|---|---|---|
| Walking | 99.7% | 0.3% | 0% |
| Waving | 3.1% | 96.0% | 0.9% |
| Standing | 0% | 6.2% | 93.8% |

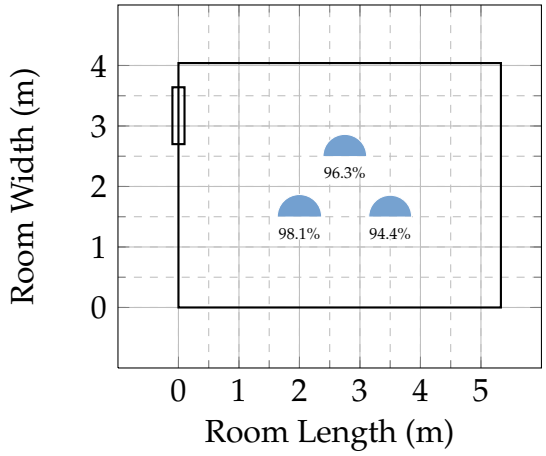
(a) Confusion matrix for R2S2

| Truth \ RFHAR | Walking | Waving | Standing |
|---|---|---|---|
| Walking | 99.1% | 0.9% | 0% |
| Waving | 5.1% | 74.0% | 20.9% |
| Standing | 0.3% | 4.9% | 94.8% |

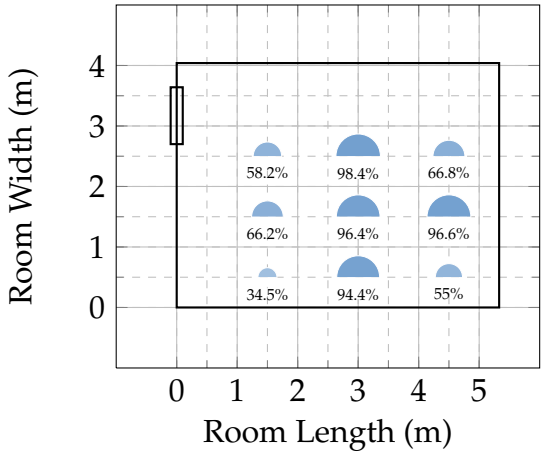
(b) Confusion matrix for R2S3A

Figure 61.: Confusions in R2S2 and R2S3A.

From the confusion matrices we observed that the recognition of the "Walking" activity was achieved in both campaigns with comparable true positive rate (TPR) of nearly 99%. This was also indicated by the confusions with "Walking" which were negligible in both confusion matrices. The same applies for the TPR for "Standing". For both campaigns this was very similar with 93.8% or 94.8%. Clearly, the strong performance difference originates from the high number of confusions of "Waving" with "Standing" in the R2S3A campaign. This figure increased by 19% points in R2S3A. We suspected that the reason for this difference is the greater number of activity locations considered in R2S3A.




(a) R2S2: TPR per location for "Waving"




(b) R2S3A: TPR per location for "Waving"

Figure 62.: True positive rate in R2S2 and R2S3A across all measured locations for the activity "Standing and Waving".

Figure 62 shows the true positive rate for each location at which the activity "Standing and Waving" was conducted in. Looking at the figure we find that R2S2 achieved above 94% TPR for each location. R2S3A also had multiple locations in which the recognition is above 94%. However, multiple locations further away from the room middle showed a significantly reduced performance.

### 8.7.2 *Temporal Recognition Stability*

In this section we investigate the performance of the trained RFHAR system over multiple days. Thus, we used the first recording of the R2S3B campaign as training set for RFHAR. Then we evaluated the trained system against all consecutive recordings from R2S3B. Figure 63 shows the RFHAR performance on the R2S3B dataset. We observed that the accuracy for the measurements on day 1 were very

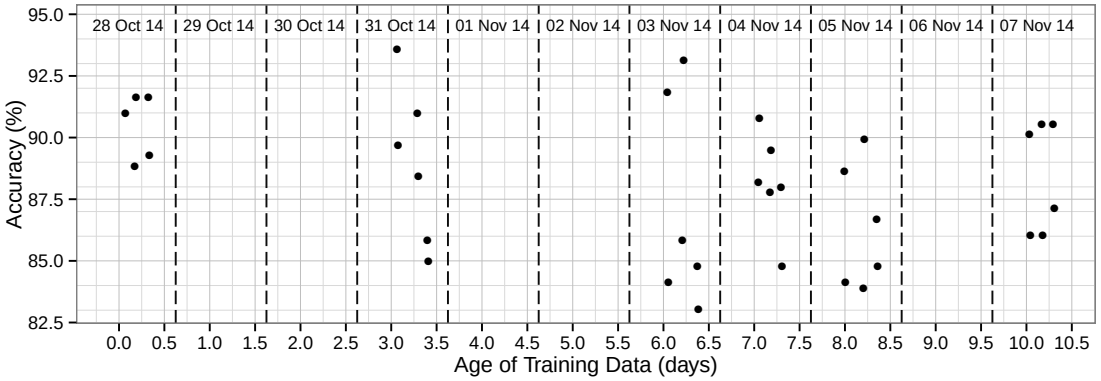



Figure 63.: RFHAR accuracy vs age of training data (R2S3B). Each dot represents the result of a single RFHAR evaluation against a recording. Broken lines indicate the beginning of a new day. The actual dates are given in the top of the figure.

closely clustered (88%-92.5%). In the following days the performance varied more strongly (83%-94%). The last day had a standard deviation similar to the first day but a lower average performance of approximately 88%. This performance is similar to the performance reported for the individual days (Sec. 8.7.1). Thus, we do not find an evidence for a correlation between recognition performance and training data. We conduct the same analysis for the R1S1 dataset which has been

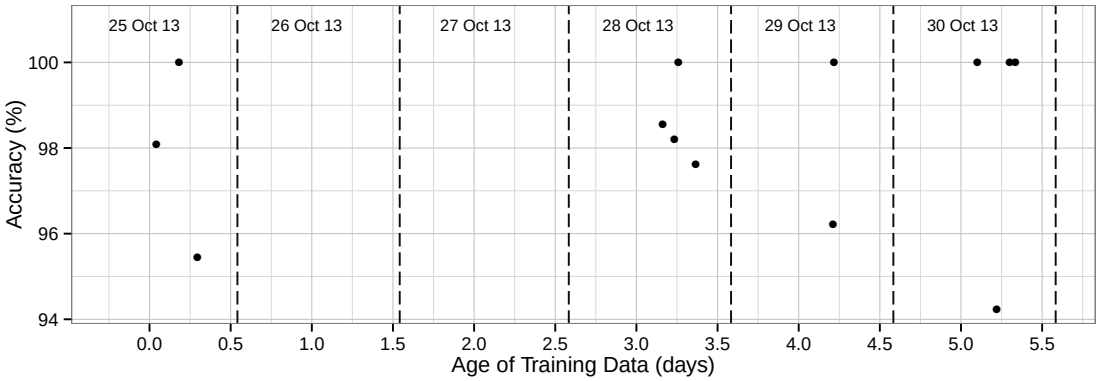



Figure 64.: RFHAR accuracy vs age of training data using R1S1 (Opened Door).

recorded over four days only. Figure 64 shows the results. Again the performance lay within the bounds (94%-100%) of the single day evaluations. We conclude that

for the investigated campaigns we cannot find a direct relationship between age of training data and accuracy.

### 8.7.3 *Influence of Training Dataset Length*

In this section we evaluate how training data length affects recognition accuracy. Training dataset length is important as training of a system is cumbersome and reducing training time would contribute to the practicability of RFHAR. On the other hand, increasing training dataset length could improve recognition accuracy. Thus, we also investigated if longer training data is helpful. For the evaluation we select the R2S3A campaign as it contains the most datasets for a single day. R2S3A contains a total number of nine recordings. We create two new subsets of data from the whole campaign: the first four recordings on this day are used as training set, the last five recordings are used for testing. The decision of taking the first four sets as training sets was based on the assumption that a user is much more likely to make the effort of recording training data at once instead of multiple times spread over the day. Based on these four training sets, we increased the data used for training up to the fourfold amount. For reducing training set size we took the smallest activity length (9 s) and reduced it by one second for each step. Then we adapted all other activities such that the prevalence was the same for each class (stratification). Figure 65 shows the results. The default training set length is 8.4 minutes. We observed a weak trend showing average accuracy increases with increasing training data. However, the small improvement does not even out a quadrupled training time. Also note that the difference between the shortest training time 54 s (89%) and the longest period 33.6 min (90.5%) is only 1.5%. Standard deviation decreased with increasing training time as well, but was still relatively large with over 2% points.

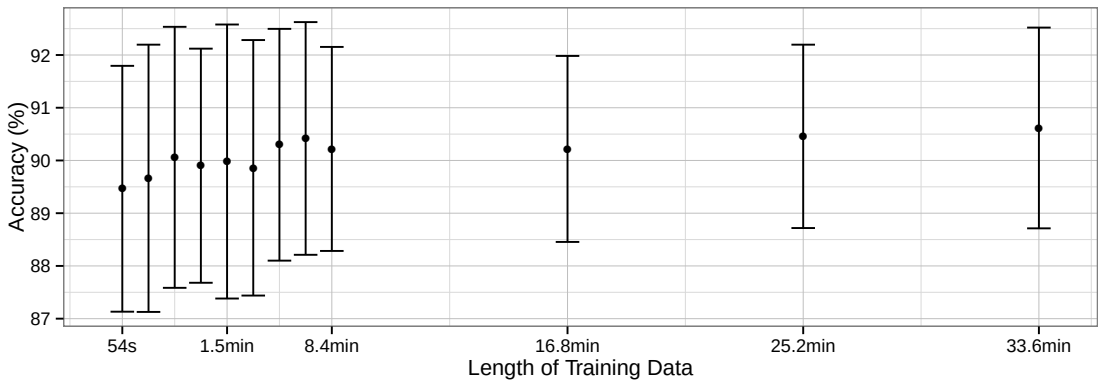



Figure 65.: Length of Training Data vs Classification Accuracy. Each dot represents the mean accuracy over the five test sets. Whiskers denote the standard deviation.

The data suggests that increasing the amount of training data has only limited impact on achieved accuracy. On the other hand, training the system with a

much shorter (and more practical) amount of data will not impair recognition performance.

### 8.7.4 *Effect of Topology Change*

In this section we consider the influence of a topology change of the Radio Sensor in respect to recognition performance. Consider the case when the system was trained and thereafter the installation was slightly changed, e.g. change of antenna orientation. Using R2S3A and R2S3B campaigns we evaluated how a changed infrastructure affects recognition performance. Since infrastructure changes affect all links associated with a changed transceiver, it was expected that this would a measurable effect on RFHAR recognition performance. R2S3A and R2S3B datasets contain the same activities and locations. But in the R2S3B campaign a single transceiver was shifted by approx. 20 cm.

Figure 66 shows the results when training RFHAR using the topology R2S3A and testing against activities conducted in the slightly changed topology R2S3B Day 1. For reference, the figure also shows the performance when testing within the same topology R2S3A.

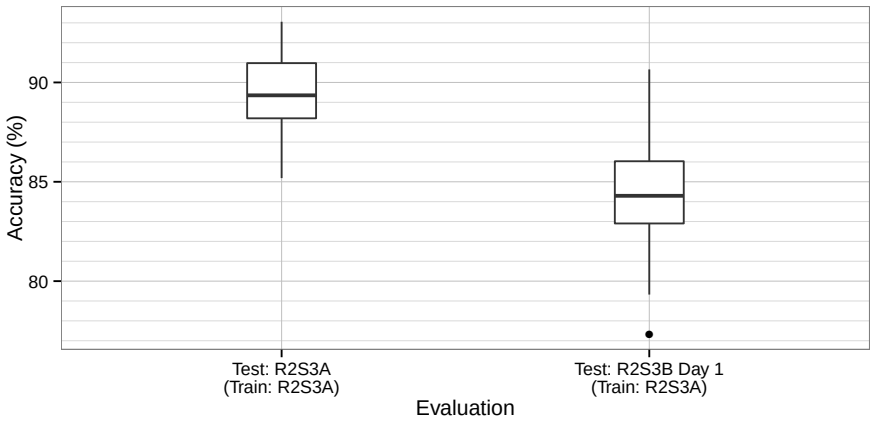



Figure 66.: RFHAR performance using the original topology (R2S3A) and after changing the topology (R2S3B)

The box plot shows a moderate median accuracy reduction from 89% (original) to 84% (changed topology). The inter quartile range was very similar which means that for the majority of evaluations performance was within the same bounds around the median as in the original ($\pm$1.5%). However, in a number of rare cases the system fluctuated more strongly which is reflected by an increased range: 85%-93% (original) and 77%-90.5% (changed topology). In Figure 67 the confusion matrices computed over all evaluations for R2S3A and the R2S3B show which activities were affected by the reduced performance.

Comparing the matrices we find that the true positive rate of "Walking" and "Standing and Waving" was not influenced by the topology change. However, the true positive rate of "Standing" was reduced by 20% as it was confused with

| Truth \ RFHAR | Walking | Waving | Standing |
|---|---|---|---|
| Walking | 99.1% | 0.9% | 0% |
| Waving | 5.1% | 74.0% | 20.9% |
| Standing | 0.3% | 4.9% | 94.8% |

(a) Confusion matrix for testing RFHAR against R2S3A

| Truth \ RFHAR | Walking | Waving | Standing |
|---|---|---|---|
| Walking | 100% | 0% | 0 % |
| Waving | 15.0% | 75.7% | 9.3% |
| Standing | 0.6% | 23.6% | 75.8% |

(b) Confusion matrix for testing RFHAR against R2S3B Day1

Figure 67.: Confusions in R2S3A and R2S3B when RFHAR was trained only using R2S3A.

"Waving" in one out of four times. This indicates that in the R2S3B topology "Standing" produces feature values which were similar to these of "Waving" in the original topology. Additionally, "Waving" was now confused more frequently with "Walking". This strengthens the argumentation: the changed topology increased the influence of activities, leading to the confusion with the next higher activity category.

### 8.7.5 *Impact of Environment Change*

Radio propagation is strongly affected by the surrounding environment. Environmental changes taking place between system training and system test are therefore likely to affect the recognition performance. In Ref. [73] we conducted a first evaluation showing that changed room geometry has an effect on Activity Recognition. However, estimating the impact was difficult due to the limited scale of the experiments and the employment of 10-fold cross-validation. In this work we leverage the multi day R1S1 measurement campaign. While the dataset considers only the activities "Walking" and "Standing", the activities were recorded with different door states. Selecting the first day of the R1S1 measurement campaign we trained RFHAR either using the open or closed door recordings and tested against the other case. For comparison, we also present the performance when training and testing on recordings with the same door state. The box plot in Figure 68 shows the results of the investigation.

The investigation reveals two findings: 1) The achieved accuracy strongly depends on the test set and much less on the training set. 2) When the room door was closed, the achieved median accuracies were improved but the standard deviation was also increased considerably. However, comparing result 2) to the performance on R1S1 in Figure 60 shows that this finding is likely by chance. In contrast, finding 1) indicates that for cat 1 and cat 3 activities the RFHAR recognition was not affected by the changed door state. This is an impressive result demonstrating the robustness of the system. Particularly, as the open door covers one of the transceivers. However, "Standing" and "Walking" are very different regarding their characteristic footprint in the RSS. In contrast, cat 2 activities, share similarities with both of these activities. Thus, we expect that performance would be degraded if a cat 2 activity, such as "Waving" would have been included in this investigation.

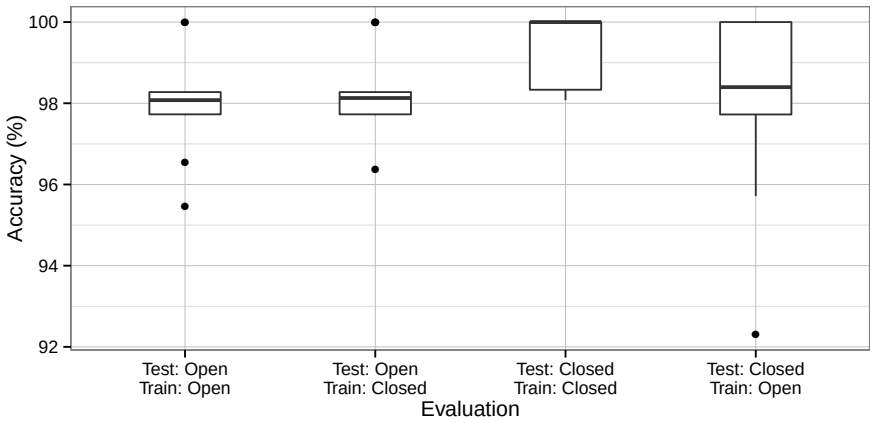



Figure 68.: RFHAR performance when the room door is open or closed during training or test

## 8.8 DISCUSSION

### 8.8.1 *Activity Categorization*

In this chapter, we classified human activities into three different categories based on their impact in the received signal strength: dynamic with changing location, dynamic with static location, and static with static location. While not explicitly stated, previous work has differentiated activities into only two categories. Shi et al. [78] used this approach to determine the classification pipeline depending on the category of the activity (dynamic/static). Similarly, Wang et al. [90] first determined the activity type into "Walking" and "in-place" and then proceeded with the actual recognition of the in-place activity. Thus, considering a third category might allow improved recognition systems. In accordance to this assumption, our evaluation showed that the three category examples "Walking", "Standing" and "Waving" were mostly distinguished with high accuracy. However, in some testbed locations "Waving" was confused with "Standing" when it did not induce strong enough fluctuations in the RSS. Thus, the difference between categories may dissolve based on the topology and measurement quantity used. However, given sufficient wireless coverage a resting in-place activity (e.g. "Standing", "Sleeping") and an in-place activity with stronger motion ("Writing", "Typing") will likely show different characteristics in fluctuation frequency and magnitude.

### 8.8.2 *Optimization*

Optimization was conducted step-wise for each RFHAR module using a single multi day measurement campaign. In the evaluation we discovered that this campaign achieved very high accuracies due to the fact that the tested locations were relatively close together and well covered by the wireless links. This also showed in the optimization, e.g. during feature evaluation all features delivered

very good results. Using a more challenging measurement campaign might have resulted in a different set of configuration parameters and possibly better results in the evaluation. Results of RFHAR could be further improved by considering the provision of an optimal feature set to the classifier [38]. Another improvement may lie in the application of alternative preprocessing filters. While Wang et al. [90] also applied low pass filtering to their CSI data, another filter worth considering is the alpha trimmed mean filter which was successfully employed by Yang et al. for RSS-based intrusion detection [97].

### 8.8.3 *Evaluation*

One of the goals of this chapter was to analyse RFHAR regarding practical Activity Recognition concerns. These include the effort of calibration and the stability of the recognition. We could show that these two aspects only slightly affected RFHAR performance. However, although we used a set of only three activities the achieved accuracy for the more elaborate measurement campaign was only 85% on average. We illustrated the meaning of this figure in the evaluation section: in some locations the system worked perfectly, while in others it was only able to differentiate between "Walking" and "other activity". It is safe to assume that if "Standing" and "Waving" can not be separated in these location other activities within each of these categories are also hardly differentiable. In this case changing the radio topology or the environment is the only way to improve the recognition. Alternatively, activities could be restricted to certain locations where they are well recognized. In their paper, Wang et al. [90] reported impressive recognition rates (>95%) for a larger number of activities over multiple days. However, activities are executed in specific locations only and it was demonstrated previously that CSI for localization achieves very good results [1]. On the other hand, they evaluated their system over the area of an apartment which is impressive as they only employed up to four links. This indicates that CSI is possibly more sensitive to environmental changes outside the proximity of the LoS than RSS. Another reason for their very good results, may be the much higher power used in WiFi (typ. 100 mW) compared to IEEE 802.15.4 (typ. 1 mW). As we have shown in Chapter 4, power has a strong effect on recognition performance as it changes coverage. Lastly, as an additional evaluation investigation it would be interesting to conduct a more in-depth study of the influence between activity and location. If a model or a transformation can be found which allows to estimate the influence of an activity on RSS in a new location based on a recorded location, then training effort could be strongly reduced.

## 8.9 CONCLUSION

In this chapter we presented RFHAR, a device-free, radio-based Human Activity Recognition system. We further provided a categorization of ADL with respect to device-free sensing based on theoretical considerations and an a priori study.

RFHAR can discriminate three ADL of different categories with 85% accuracy on average even when reducing training time to a minimum (1 min) or when testing the system up to 10 days after training. Recognition is also only slightly affected when changing the sensor topology. Hence, this work is the first to demonstrate long-term RSS-based Activity Recognition on independent datasets. These results were achieved through the development of a novel frequency fusion mechanism and a novel feature. Future research directions should include the extension of the system to include additional activities, to integrate multiple features and alternative preprocessing steps and the investigation of location-independent activity characteristics, among others. With the successful implementation of RFHAR we consider the requirement of a Activity Recognition Inference System for a practical Activity Recognition system as described in Chapter 5 as satisfied.

# 9 Discussion

In this section we discuss a number of aspects of this thesis on a more general level. Discussions of specific topics are provided in the corresponding chapters of this work.

## 9.1 Software design pattern for inference systems

In Chapter 5 we derived a software design pattern for radio-based Inference Systems based on previous work from classical Activity Recognition and device-free context recognition. For all three presented Inference Systems (presence, discrimination and activity) the pattern was applied successfully. It is important to highlight that the pattern allows to use external means of model generation/training - one major difference to the classical machine learning processing chain. An example of a system exploiting this aspect is the WiDisc system in Chapter 7. In order to investigate the general validity of the pattern, we explore some recent device-free, radio-based context recognition systems.

The WiGest system [2] is a sophisticated gesture recognition system which does not require calibration. However, gesture identification i.e. inference is performed using sensor data preprocessed in multiple stages and a *predefined* set of gesture families. If interpreting the definition of this gesture set as model generation, which is valid as (preprocessed) sensor data is assigned a groundtruth-based meaning, the WiGest system architecture is compatible to the presented design pattern. The BreathTaking system by Patwari et al. [60] employs a model developed offline for breathing rate estimation. In contrast, to the Inference Systems developed in this work and WiGest, its output is not categorical but ordinal (breathing rate). Nevertheless, it is still within the bounds of the proposed architecture. Finally, the RASID presence detection system [45] diverges from the other approaches as it continuously updates its model created in offline calibration during online phase to cope with environmental changes. This design is also covered by the presented architectural pattern as the *Inference* module may access the *Model* for reading but also for writing.

Hence, we conclude that the proposed design pattern reflects most prominent approaches for device-free, radio-based context recognition systems. While the provided analysis is not exhaustive and it is likely that further designs exist to approach device-free context recognition challenges the pattern provides a

helpful guide to identify the various possibilities for implementing novel Inference Systems and improving existing systems.

## 9.2 HOLISTIC SYSTEM ARCHITECTURE FOR PRACTICAL ACTIVITY RECOGNITION

In Chapter 5 we proposed a holistic system architecture combining the Radio Sensor and multiple inference subsystems to enable practical Activity Recognition. We have developed and successfully evaluated the proposed subsystems in the preceding Chapters and Appendix A. Here we discuss the combined holistic system. Figure 69 shows the conflated architecture with the most important

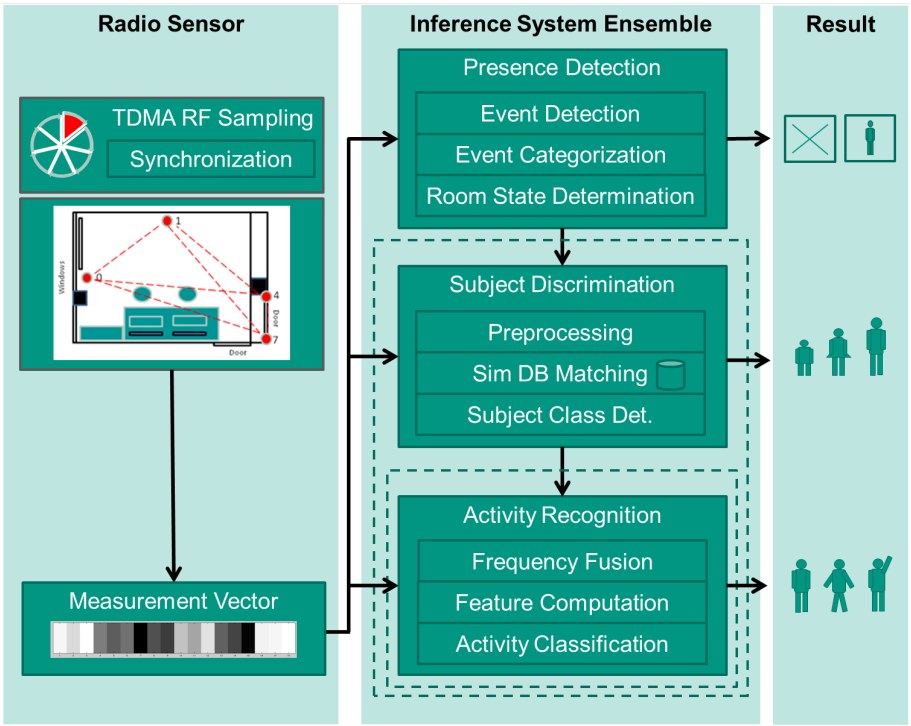



Figure 69.: Holistic System Architecture with Subsystems

modules of each subsystem. In order to estimate the overall system performance we summarize the accuracies for each Inference System in Table 16. Note that for RFHAR we use the median accuracy achieved on the R2S3A datasets.

| | **Presence Detection** LoS-Cross | **Subject Discrimination** WiDisc | **Activity Recognition** RFHAR |
|---|---|---|---|
| Accuracy (%) | 92% | 67% | 89% |

Table 16.: Achieved Accuracies of the different Inference Systems

The holistic system can be operated in three modes as follows:

1. Direct Mode
   Input: Radio Sensor measurements
   Output: Presence, Subject Class, Activity as discrete results

2. Presence-based Mode
   Input: Radio Sensor measurements, Presence Detection result
   Output: Not Present or Subject and Activity

3. Presence and Subject-based Mode
   Input: Radio Sensor Measurements, Presence Detection result, Subject Class
   Output: Not Present or Subject and Activity (based on subject profile)

In Direct Mode the performance of the holistic system would be comprised of the performance of every individual system. Thus, it would be identical to the results shown in Table 16. In Presence-based Mode the holistic system performance would be comprised of the accuracies of WiDisc and RFHAR weighted by the LoS-Cross result. Hence, performance would be 62% and 82% for the Subject Discrimination and Activity Recognition, respectively. In Presence and Subject-based Mode the system performance would be the accuracy of RFHAR weighted by the LoS-Cross and WiDisc results (54%) and the WiDisc result weighted by the presence information (62%) as before. However, at this point RFHAR does not use subject-specific calibration. Hence, the Presence and Subject-based Mode provided no advantages. Also note that the given numbers are just rough estimations as the systems were evaluated in different testbeds and parameter settings. Using the Direct Mode is also not plausible as the results depend on each other, e.g. Subject Discrimination and Activity Recognition systems require that a person is present. Therefore, for the holistic system the Presence-based operation mode is advisable. Thus, we can estimate the overall system performance with 62% and 82% accuracy, for Subject Discrimination and Activity Recognition, respectively. The fact that Subject Discrimination requires a subject to stand still for 5 s and that the number of currently recognizable activities is very limited, highlights the need for further research.

Comparing the estimated overall system performance to other research is difficult as no comparable device-free, radio-based system has been previously presented. We find that the closest work to ours is the work of Wang et al. [90] as they considered Activity Recognition and subject presence and conducted their evaluation using independent data sets. They achieved 90-97% depending on the used test bed. However, they primarily considered activities that differed per location which likely improves recognition performance strongly. In addition, they relied on channel state information which provides more radio signal information but is currently not available in existing infrastructures. We have discussed their results in more detail in Chapter 8.

Nevertheless, the combined results make the employment in real world environments difficult as the system still makes too many mistakes. With the selected operation mode the execution is cascaded, i.e. the performance of the first Inference System has tremendous influence on the performance of the subsequent system. Hence, improving the performance of the Presence Detection is most important. One strength of the architecture, which is helpful in this regard is the modularization of the Inference Systems with respect to the investigated context.

For instance, Presence Detection could also be improved by replacing LoS-Cross with a system leveraging a different kind of available presence sensor e.g. PIR, light switch, etc.

Nevertheless, as we demonstrated that each of the requirements defined in Chapter 5 can be addressed using RSS, Thesis 3 has been validated.

## 9.3 COMPARISON TO OTHER SENSOR SYSTEMS

A strong motivation of this work were the disadvantages present in current sensor systems hindering their employment in the Smart Home. Both were defined in the introduction chapter.

The first disadvantage was the installation effort. Appendix A explains how the Radio Sensor was developed such that it provided an ideal basis for our research. However, with these changes we diverged in some aspects from standard IEEE 802.15.4 functionality. Nevertheless, as discussed in the same chapter it is technically possible that the original services and the Radio Sensor functionality can be realized on the same device – probably only requiring a firmware update. Assuming that sufficient smart devices are distributed across the home, the system can be employed without additional installation effort, which was the initial disadvantage of other systems.

The second disadvantage we identified in existing solutions was the comfort of use. In the introduction we defined this as the user awareness level of the system. For instance, if the system requires additional actions which the user would normally not do, e.g. attaching the system to the body after getting up, comfort of use is reduced. It also decreases the likelihood of using the system as the user may simply forget to conduct these activation actions. Unfortunately, although device-free, radio-based Activity Recognition has great potential in this regard, comfort of use is also an issue. Although we have shown in Chapter 8 that the recognition was stable over a time period of 10 days, we can assume that after some larger environmental change (e.g. new furniture) re-calibration becomes necessary. In addition, the achieved recognition results still cannot compete with other systems. But good recognition performance is especially important to realize implicit interaction. As there is typically no immediate action in reaction to a user activity, the user will not repeat his activity to trigger the action. Hence, there is also no chance of improving results by repetition. In contrast, this is possible, when explicitly controlling a system using gestures as done for instance in Ref. [2]. Hence, achieved recognition accuracy must be further improved and calibration should be replaced with alternatives.

## 9.4 OPEN CHALLENGES FOR DEVICE-FREE, RADIO-BASED HAR

### 9.4.1 *Estimation of Radio Sensor Coverage*

It has been shown over the course of this work that accuracy is largely influenced by the coverage of the monitored area (cf. Chapter 4 and Chapter 8). Of course, other parameters such as inter-subject differences also influence accuracy. However, if no measurable wireless signals penetrate the area of interest in which an activity is conducted it cannot be recognized. The straight forward way to approach this challenge is deploying as many radio transceivers as possible [91]. The advantage is that the radio transceivers are visible and hence coverage can be roughly estimated. Another approach is to tune parameters in the Radio Sensor such that the coverage area is changed, e.g. frequencies, transmission power, transceiver sensitivity and resolution. However, without suited predictive models and corresponding automated tools this approach lacks the ability to ensure coverage. This is also true for the other two possibilities: measure additional signal characteristics which may be more sensitive to subtle changes in areas with limited coverage [90] or utilize ambient radio signals [78]. Hence, currently to ensure best recognition performance it is advisable to deploy as many transceivers as possible. Depending on the type of activities the actual topology should also be adapted accordingly [73].

### 9.4.2 *Automatic Calibration*

Calibration enables recognition in a new sensor system installation with unknown parameters. Depending on the underlying algorithm the process of calibration may only determine a few variables (e.g. threshold estimation) or rebuild the entire model involving the creation of a pattern database and the assignment of the actual meaning of the instances (e.g. classifier training using groundtruth). The presented Inference Systems require manual calibration with the exception of WiDisc which replaces manual calibration by simulation. However, generating the necessary 3D models also requires some effort. As discussed earlier, ideal comfort of use means calibration does not require user effort. Multiple approaches to automatic calibration can be imagined: In transfer learning [57] a trained system is transfered to a different feature space than it was originally trained with. In other words, with transfer learning an Activity Recognition sensor system trained in the lab may be transferred to a Smart Home. The challenge here lies in discovering a suitable mapping function without manual intervention. A second possibility is the opportunistic use of existing sensors which may provide the groundtruth to automatically conduct the calibration (cf. [47]). For instance, Radio Sensor measurements may be annotated with the information “subject present” when a light switch in a room is activated. While at night the light switch may be sufficient to determine subject presence, the automatically trained radio-based system may now detect presence during the day. Challenges associated with this approach are

the access of existing sensor data and the inference of activities not measurable with other sensor systems. A third more complex approach lies in the estimation of unknown parameters using the sensor system itself. Starting from very limited groundtruth such as the distance between transceivers, this may be done by performing continuous calibration measurements iterating through all adaptable Radio Sensor parameters while optimizing an internal world model. If this world model has been created successfully a user model could be introduced. Using wave propagation algorithms the system could then determine user presence, activities and other contexts based on Radio Sensor measurements.

### 9.4.3 *Practical Activity Recognition for multiple Subjects*

We have excluded multi user Activity Recognition from this work because we wanted to investigate the feasibility of the approach before raising the complexity. Indeed for most of the presented Inference Systems multi subject-based evaluation increases the evaluation effort tremendously. The reason is the underlying algorithmic concept which is based on supervised machine learning. Hence, patterns for each activity combination of each subject in each position must be provided as training data to enable recognition. For instance, the investigation of five activities by two individuals by Sigg et al. [84] required to record all 25 combinations as a training basis for the system. While it is impressive that researchers went to such great lengths to proof feasibility, it is unrealistic to require the same from a user. The challenge here is clearly to develop approaches which allow multi user Activity Recognition without this immense overhead. Models such as those suggested for calibration may provide a solution. Alternatively, device-free localization research successfully demonstrated multi subject localization, e.g. [95] which may offer approaches for multi subject Activity Recognition. In addition, classical Activity Recognition research also offers a corpus of research considering multiple subjects. Also note that the developed Inference Systems are all room level systems. Hence, if the goal is to recognize single user activities in multiple rooms, the presented system approach may suffice.

# 10 Conclusion and Outlook

In this thesis we demonstrated that the recognition of single user activities using COTS Smart Home wireless transceivers is feasible. From the experimental investigation in Chapter 3 we can further summarize the following findings:

- Device-free, radio-based Activity Recognition can provide the same performance as device-bound recognition (in 10-fold cross validation).
- Depending on the coverage also very small motions can be detected and identified in radio signals.
- Combining Activity Recognition and Presence Detection in a single classifier reduces performance.

As a next step we showed that recognition performance is influenced by three main aspects (Radio Sensor, environment, subject) and their parameters to different extent. From the corresponding investigation in Chapter 4, the following lessons were learned:

- The impact of investigated parameters on recognition performance decreases as follows: size of environment, transmission power, frequency spacing, number of used frequencies and link length.
- Dynamically adapting Radio Sensor parameters typically leads to improved recognition. Examples include carrier frequency and transmission power.
- Location and context specific radio coverage is needed for optimal recognition results.
- When using short wireless links even slight execution differences in the same activity can be identified. Longer wireless links are helpful to distinguish between activities as intra-activity differences converge. If links are too long (in our case >4 m) the influence of activities fades and they will be hardly separable.
- Larger environments typically result in a reduced recognition rate as radio signals need to travel further until being reflected back into the area of interest.
- Measuring at different frequencies concurrently, increases the information gathered from the environment. Spacing frequencies further apart increases the amount of captured information.

To implement a practical device-free, radio-based Activity Recognition three challenges were identified (cf. Chapter 5) and successfully addressed in this thesis: Presence Detection, Subject Discrimination and Activity Recognition.

In Chapter 6, a novel Presence Detection system was developed and evaluated. It has an accuracy of 92% and employs a specific transceiver topology to detect even a non-moving subject. Lessons learned from this investigation were as follows:

- The range feature is better suited than standard deviation to capture LoS crossings; combining the two features may improve recognition performance further.
- Very short sliding windows (0.25 s) are better suited to capture LoS crossings than longer (>0.5 s) or very short windows (0.125 s). This also depends on the speed at which the crossing takes place and on the length of the link (in our case: 1.20 m).
- Approximating the correct spatial expansion of such short indoor links is non-trivial as they may be distorted by the immediate surroundings.
- Motion in close proximity to the LoS can cause fluctuations of comparable magnitude to a LoS crossing.

The Subject Discrimination system described in Chapter 7, can differentiate three subject classes with an accuracy of 67% using 3D modelling and radio propagation simulation to circumvent cumbersome manual calibration. The following insights were gained during development and evaluation of this system:

- When employing simulation as part of the system the most probable source of error is divergence of simulation and measurements.
- In our case the simulation result was noisier than the measurement. Changes of different 3D model parameters did not yield an improved simulation result. Consequently, it seems that the prediction of wireless indoor propagation has to be improved to make simulation for device-free recognition even more useful.
- The difference between simulation and measurements was reduced when averaging the three measured/simulated frequencies per link.
- The relationship between a subject's dimension and the measured signal is non-trivial due to complexity of indoor radio propagation (i.e. tall != strong attenuation).

The Activity Recognition system described in Chapter 8, employs a novel feature and several new preprocessing steps. In an experimental evaluation spanning ten days the system showed a stable average recognition accuracy of around 85% for three different activities conducted in two different orientations and in nine distinct locations. Further lessons were:

- Activities can be grouped in three categories based on their impact on the radio signal. The impact is determined by the amount of motion intrinsic to the activity, the dimension of the moved object and the subject's location.

- Activities which only have slight motion components typically require a better coverage for recognition.
- Activities with stronger impact, e.g. "Walking", usually require minimal topology and transceiver density.
- If coverage is sufficient, training data can be minimal (i.e. a single repetition of the activity in the specific place and orientation is sufficient).
- One possibility to further reduce training effort, could lie in algorithms which can estimate the signal fluctuation in new locations/orientations through suited models.
- Minor changes in the radio sensor topology have considerable influence on the recognition performance.
- Fusing measurements from different carrier frequencies of the same link improves Activity Recognition performance.

The combination of Presence Detection, Subject Discrimination and Activity Recognition into a holistic hierarchical architecture provides the first personalized Activity Recognition on room level using this sensor type. It presents a first ambitious attempt to provide all these information at once from ubiquitous radio signal characteristics. However, exciting challenges remain to make device-free, radio-based context recognition a genuine option for implicit interaction in the Smart Home. Such challenges include multi subject recognition, automatic calibration, context-specific coverage estimation and the exploration of novel contexts which benefit from the characteristics of radio, among others.

# A RADIO SENSOR

This chapter describes the technical implementation of the Radio Sensor. Hence, it presents the practical basis for the radio channel measurements for all experiments in this dissertation.

The chapter is structured as follows. First, we review related work regarding wireless technologies and existing sampling algorithms. Then we conduct an analysis of the requirements and existing research. Based on the conclusions we describe our implementation and validation of the Radio Sensor in the next section. The chapter is closed with a discussion and a conclusion.

## A.1 INTRODUCTION

In order to conduct investigations for device-free, radio-based Activity Recognition a sensor is needed which provides periodic radio channel measurements. Throughout this dissertation we refer to this sensor as Radio Sensor. Recapitulating the definition in Chapter 4, a Radio Sensor is a configuration of a single or multiple hardware entities receiving radio signals. Radio Sensor measurements are based on these received signals. A Radio Sensor actively transmitting radio signals is called *active*, while it is called *passive* when measurements are generated only by listening in on ambient signals. Using a passive Radio Sensor increases uncertainty as the signal sources are not in control of the sensor. For this reason we developed an *active* Radio Sensor. As it is the most common radio signal characteristic, the developed Radio Sensor measures the received signal strength (RSS). RSS is a measure which quantifies the power of the superimposed electromagnetic waves reaching the receiver antenna.

Based on the vision of this work and the parameter investigation in Chapter 4, we defined the following requirements for the sensor:

- Smart Home compatibility: The sensor should operate using a wireless technology common in the Smart Home.
- Sampling rate tuned for Activity Recognition: The sensor should sample the wireless channel fast enough for every human activity.
- Transmission Power: The sensor should use the highest transmission power available in the transceiver.

- Frequency Diversity: The sensor should allow to employ multiple frequencies.
- Frequency Spacing: The sensor should operate on frequencies spaced as far apart as possible.
- Topology: The sensor should incorporate as many transceivers as possible.

In order to satisfy these requirements we first reviewed Smart Home wireless communication standards. It is shown that the most suited standard is the IEEE 802.15.4 standard and that a novel sampling algorithm for RSS measurement must be developed. Hence, we implemented our own Radio Sensor directly in the firmware of the 2.4 GHz IEEE 802.15.4 NXP Jennic 5139 transceivers. The implemented time division multiplex algorithm provides measurements with 40 Hz for wireless links between up to 10 transceivers or 4 transceivers when measuring on a single or three carrier frequencies, respectively.

## A.2 BACKGROUND AND RELATED WORK

### A.2.1 *Smart Home Wireless Communication Technologies*

While analysts predict the bright future of the Smart Home market [4], *the* Smart Home wireless communication technology has yet to be identified. Today, many standardized technologies are available [17]. Depending on the type of device these include WiFi, Bluetooth, IEEE 802.15.4 (which is the foundation for the ZigBee stack) and many others. However, we may safely assume that the ones mentioned are the most popular wireless Smart Home technologies today. WiFi is near ubiquitous due to the comfortable provision of internet access which gave it a massive rise in recent years. Bluetooth on the other hand, is pervasive due to the large spread of mobile phones and other wearables which typically carry Bluetooth functionality. Interestingly, IEEE 802.15.4 which was build as a low power solution for Internet of Things and Smart Home systems is currently only available in few products such as remote controlled lighting. Nevertheless, new efforts such as the Thread Group [1] which is powered by a number of well known technology companies may finally bring ZigBee into the Smart Home. Thus, the next paragraph briefly describes these technologies. Descriptions are in parts based on Ref. [24].

Wireless Local Area Networks (IEEE 802.11 WiFi [19]) are optimized for range (typically 30-100m) and high data rates (up to multiple Gbit/s with the newest standards). WiFi is an extension of Local Area Networks (LANs). Thus, once a device becomes part of the network it is treated identical to other devices in the network. They are typically operated in infrastructure mode in which an access point provides mobile units access to other networks.

1 http://threadgroup.org, retrieved: 23 Feb 2015

On the other hand, Bluetooth [80] and IEEE 802.15.4 [35] are Wireless Personal Area Networks (WPANs). These were designed with the focus of providing power-efficient wireless communication in the direct environment of the user without the need for an infrastructure. Bluetooth was originally developed for data rates of 1-3 Mbit/s and a range of up to 10 m and for services such as wireless headsets. In contrast, IEEE 802.15.4 was developed for even slower bit rates of up to 250 Kbit/s but a range of 10-100 m.

Typical application cases for IEEE 802.15.4 are the transmission of control commands or the retrieval of sensor values from some appliances. Originally, IEEE 802.15.4 was designed so that the transceivers can be very simple and therefore only consume a fraction of the energy required for Bluetooth or WiFi. But Bluetooth 4.0 Low Energy has closed the consumption gap and directly targets applications which used to be a perfect fit for IEEE 802.15.4. However, a very specific difference between both protocols still lies in the connection scheme. While Bluetooth requires a dedicated connection between devices, IEEE 802.15.4 allows broadcast communication facilitating the simple construction of fully meshed networks.

### A.2.2 *Algorithms for Sampling Received Signal Strength*

A number of RSS sampling algorithms for commodity hardware have been developed in device-free localization research. Typically, nodes are utilized using a time division multiple access scheme (TDMA). However, TDMA requires time synchronization in order to ensure that nodes do not interfere when sending. As the transceivers typically carry low-quality crystal oscillators, synchronization is the most important purpose of these algorithms. Patwari and Wilson use a token passing protocol to implement their sampling algorithm [91] called SPIN on a IEEE 802.15.4 TelosB motes with TinyOS. In this protocol the node that currently has the token sends a broadcast packet. The packet is received by all other nodes in the network and received signal strengths are calculated. Then the next node in sequence sends the broadcast and so on. If a packet is not received from the previous node, the next node sends a packet after a timeout. However, they reported no maximal TDMA slot length for a single node. In contrast, Kaltiokallio and Bocca [40] implement a continuous synchronization on the Sensinode U100 Micro.2420. The synchronization is implemented by adapting to the clock of a sink node which provides a global network time in a specific synchronization phase. Therewith the protocol has a constant error of $\pm 5\,\mu$s. The authors use the synchronization for power management. The reported TDMA maximal slot length (sending or receiving) for a single node is 4.05 ms.

## A.3 ANALYSIS

### A.3.1 *Requirements*

In the introduction we have defined a number of requirements for the Radio Sensor. We will go over each of these requirements and specify them more closely in this subsection.

The first requirement was a sensor sample rate sufficiently high to capture all human activities. Following Bouten et al. [13] who investigated human activities using accelerometers, these may include frequencies of up to 20 Hz. Hence, according to the Nyquist theorem we require the Radio Sensor to provide a sample rate of 40 Hz or 40 packets per second for each wireless link. This is a very demanding requirement for a sampling algorithm on low cost, commodity hardware. Since the SPIN algorithm by Wilson and Patwari [91] was primarily designed to flexibly and robustly integrate a high number of transceivers it has no mechanisms to guarantee channel measurements at high speeds and deterministic frequency. Hence, it is not ideal to fulfill this requirement. In contrast, the algorithm of Kaltiokallio and Bocca [40] has a maximal send slot time per transceiver. This is guaranteed through continuous synchronization. In other words, it allows to take channel measurements in fixed periodic intervals. The maximal slot time is 4.05 ms. Hence, using this algorithm we could operate $\frac{1000\,\text{ms}}{40\,\text{Hz}\times 4.05\,\text{ms}} = 6.17$ transceivers on a single frequency measuring with 40 Hz between all wireless links.

The second requirement was to employ technologies which are likely to appear in the Smart Home. In the previous section we presented three candidate technologies: WiFi, Bluetooth and IEEE 802.15.4. However, the only technology allowing the mentioned high packet rate as it allows broadcast packets on the MAC layer while at the same time being highly configurable is IEEE 802.15.4.

The third and fourth requirement are both related to the possibility of sampling RSS on different frequencies. This is possible using IEEE 802.15.4 hardware. It further suggests to use IEEE 802.15.4 in the 2.4 GHz band which offers 16 2 MHz-wide radio channels ranging from 2405 MHz up to 2480 MHz.

Lastly, it was required that as many transceivers as possible should be included in the Radio Sensor. The purpose of this requirement is clear. Increasing number of transceivers increases the number of wireless links and consecutively the coverage of the monitored area. Which may in turn lead to improved recognition. In an earlier parameter investigation [73], we showed that for an office room of 4.0 m x 5.0 m accuracy improves for increasing number of transceivers. However, the difference between using 4 or more transceivers was negligible. While this is surely not the case for larger rooms, we define four nodes (6 wireless links) as the minimal requirement. Unfortunately, requiring as many radio nodes as possible is opposed to the requirement of employing as many frequencies as possible. As the algorithm is likely to be implemented in a TDMA approach, every slot can be either filled with a additional frequency measurement or a transceiver. In effect, the number of available TDMA slots when using $n$ frequencies is the number

of overall slots divided by $n$. A Radio Sensor using the sampling algorithm of Kaltiokallio and Bocca [40] has 6 TDMA slots available. Let us assume that this algorithm could be implemented with frequency diversity without additional latency. Then using two or three carrier frequencies would allow to employ three transceivers or two transceivers, respectively. Thus, with this algorithm we cannot achieve the minimal goal of four transceivers in combination with frequency diversity. For this reason, we chose to develop our own Radio Sensor algorithm.

### A.3.2 *Implementation Approach*

Based on the defined requirements we implemented the Radio Sensor in four steps:

1. Select: Choose an appropriate IEEE 802.15.4 transceiver.

2. Benchmark: Investigate the device for its sending/receiving latency.

3. Implement: Analyze the clock drift on the selected transceiver and implement a protocol which compensates the drift in order to achieve a constant sampling rate.

4. Validate: Test algorithm with the full set of transceivers.

### A.3.3 *Summary*

In this section we specified the Radio Sensor requirements more closely based on related work. The specification is summarized in Table 17. An important conclusion from this section is that a novel Radio Sensor algorithm must be implemented to suit the needs for device-free, radio-based Activity Recognition. The implementation was performed in four steps and is described in the next section.

| **Radio Sensor Requirement** | **Value** |
|---|---|
| Type | Active |
| Measurement | RSS |
| Wireless Technology | IEEE 802.15.4 |
| Sampling Rate | 40 Hz |
| Frequency Diversity | Yes, number depends on hardware latency |
| Frequency Spacing | Maximal |
| Topology Minimum | 4 transceivers (=6 links) |
| Algorithm | New implementation |

Table 17.: Summary of Analysis

## A.4 IMPLEMENTATION AND VALIDATION

In this section we describe the selection of a IEEE 802.15.4 radio transceiver. The determination of the maximal packet rate and slot time. The implementation of the sampling algorithm and the validation of the algorithm in an experiment over a longer time period.

### A.4.1 *IEEE 802.15.4 Radio Transceivers*

In order to implement the Radio Sensor we must select an appropriate radio device. A growing number of IEEE 802.15.4 transceivers is available today[2].

We may categorize devices in two groups:

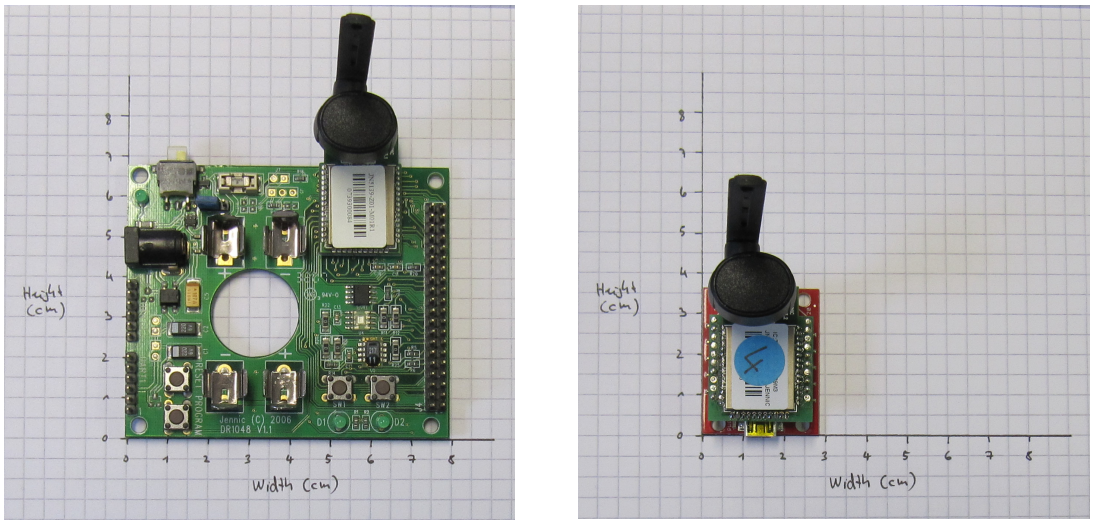
1. Rapid Prototyping: Devices which are usable with minimal configuration effort. Typically these devices abstract most of the protocol functionality into a simple end user API and provide a serial interface for configuration and data. Uploading user code on to the device is typically not possible.

2. Flexible Configuration: Devices which allow complete control of the Medium Access Control (MAC) layer, e.g. to implement a user-defined communication stack. Loading user code on these devices may be possible if they include an integrated microcontroller.

A prime example of the first category are the Digi XBee modules [36]. Examples of the second category include the Texas Instruments Chipcon modules such as the CC2531 [87], the Atmel AT86 modules e.g. the AT86RF233 [88] or the NXP Jennic [50], among others. Clearly, devices of the first category are easier to employ but much harder to constrain to specific timing requirements as required for a deterministic sampling algorithm. Hence, we consider only modules from the second category for further development.

Previously, we have used the NXP Jennic as research platform for investigating device-bound RSS effects and building domain specific wireless sensor network architectures [76, 72]. We have further ported the open source operating system Contiki to the platform and developed multiple application specific print circuit board (PCB) designs which we bundled and made accessible in the Jennisense project[3]. To leverage this corpus of knowledge we therefore chose the Jennic JN5139 as transceiver for our implementation of the Radio Sensor. The Jennic is a programmable 32 bit OpenRISC core with a 2.4 GHz IEEE 802.15.4 transceiver. It operates at 16 MHz and has 128 kB ROM and 128 kB RAM and provides a MAC level API. Alternatively, it may be operated using a ZigBee stack provided by NXP. RSS measurements range from -98 dBm to -11 dBm with a resolution of 2dBm-3dBm, resulting in 46 different signal levels. As the evaluation boards for this hardware are rather large we developed a PCB specifically optimized for

2 http://en.wikipedia.org/wiki/Comparison_of_802.15.4_radio_modules, retrieved: Feb 24, 2015
3 http://github.com/teco-kit/Jennisense/wiki, retrieved: Feb 24, 2015




(a) Devkit board (b) JBee board

Figure 70.: PCB boards for the JN5139 transceivers. Transceivers are outfitted with the Gigaant Swivel Antenna.

radio-based sensing. The board has a minimal footprint and is compatible with the popular XBee socket allowing to use it in rapid prototyping environments. Most importantly, it can be programmed with much more ease than the original evaluation board. Figure 70 shows the size comparison between the development kit board and the newly designed board named JBee. JBee schematics are openly available in the Jennisense repository. Note that the footprint may be further reduced by working directly with the transceiver chip instead of the module. However, in this case the radio circuitry has to be adapted which is a considerable effort.

### A.4.2 *Determining Maximal Slot Time*

Initially, we estimated the maximal slot time by the means of the maximal packet rate for this transceiver. For this we considered two scenarios involving only a single link i.e. two transceivers. In the first scenario one transceiver sent small IEEE 802.15.4 packets as fast as possible to the second transceiver. In the second scenario the first module sent a packet, and the other module replied after receiving this packet. Then the first module sent a packet again. This second mode was closer to the intended algorithm as it involves processing of the received packet. In both cases the received packets were either directly processed on the firmware or output to the serial to an attached PC. The Jennic modules were operated with a baud rate of 1 Mbit/s. We ran all four tests for approx. 10 min and observed no packets losses or corruptions. Table 18 shows the achieved packet rates.

We observed that the use of the serial output does not reduce the packet rate. The second observation was that the packet rate is reduced when requiring the second module to receive, parse and send a new packet. In a Radio Sensor each transceiver

| Scenario | Average Packet Rate |
|---|---|
| One Way Transmission | |
| (no serial output) | 394 Pkg/s |
| (with serial output) | 394 Pkg/s |
| Round Trip Transmission | |
| (no serial output) | 352 Pkg/s |
| (with serial output) | 352 Pkg/s |

Table 18.: Packet rates for the Jennic JN5139

must receive and process the received packets in order to determine and store the measured signal strength. Thus, we assume that the second measurement was much more realistic. Using the average packet rate we compute that a time slot would be $\frac{1000\text{ms}}{352} = 2.8\text{ms}$. However, while packet reception was nearly instant, we discovered that JN5139 has an indeterministic sending behavior. We determined the latency by measuring the time between issuing the send request in the firmware and the callback which is received when the packet is actually sent. To further ensure that the stack callback really reflects the successful sending of the packet we connected the sender and the receiver to an oscilloscope. An annotated screenshot of the measurement is shown in Figure 71.

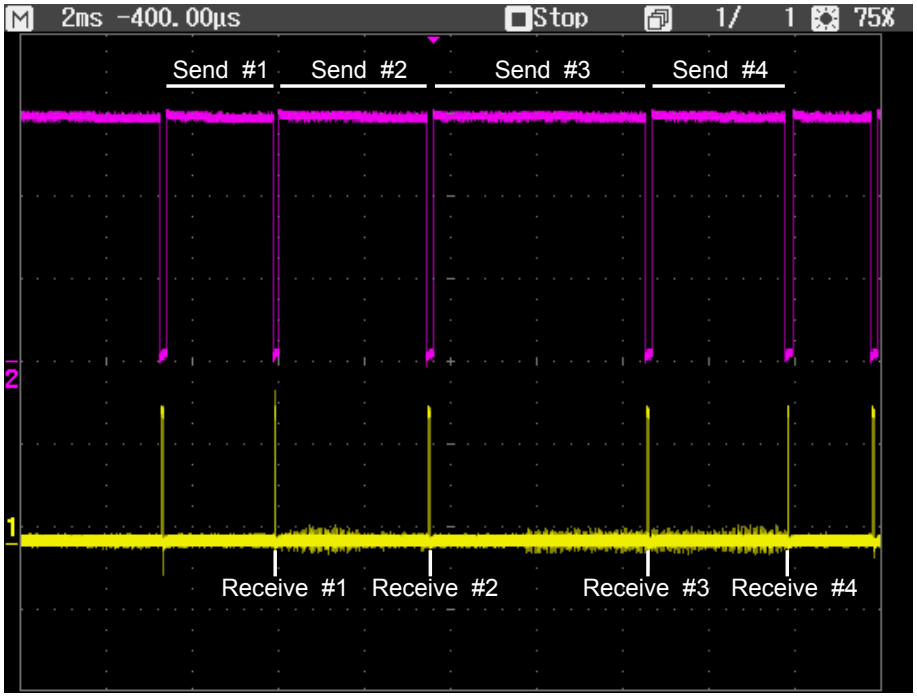



Figure 71.: Indeterministic send delay of JN5139 and the packet reception time.

In the figure, channel 2 (above) was connected to the sender. It toggled a digital pin when a packet send was requested and toggled the pin again when the stack callback informed the module that the packet had been sent. The output of channel 1 (below) was connected to the receiver. The receiver toggled a pin when the packet

was received and again when it was processed. We observed that receiving is nearly instantaneous. On the other hand, sending took between 2-5 ms. This meant time slots would be 5 ms long. Which is even longer than the slot length used by Kaltiokallio and Bocca [40]. Hence, the number of transceivers would be very limited.

By further investigation it became clear that this behavior was related to a feature of the IEEE 802.15.4 standard. IEEE 802.15.4 provides a mechanism called Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) [24]. With this mechanism the transceiver performs a clear channel assessment (CCA) of the selected channel prior to a transmission. If the channel is busy (energy above a certain threshold) it delays sending by a random time period (aka random back off). In the case of JN5139, we observed a random delay even when the channel was free. We deactivated the random back off delay. Additionally, we disabled retransmissions in the case of collisions as this would impair our sampling algorithm. Using these modifications the transmission latency became quasi deterministic at 2 ms. Adding a safety of 0.5 ms, resulted in a time slot length of 2.5 ms for the JN5139. Giving a working synchronization this allowed the use of up to $\frac{1000\,\text{ms}}{40\,\text{Hz}\times 2.5\,\text{ms}} = 10$ transceivers in the Radio Sensor while obeying the defined sampling requirements.

### A.4.3 *Synchronization Algorithm*

A Radio Sensor measurement is the current received signal strength of all wireless links between all transceivers of this sensor. To acquire this measurement, a packet originating from every module must be received by every other module. To avoid additional overhead this is typically realized using broadcast packets. Hence, for each Radio Sensor measurement a single broadcast packet must be sent by each transceiver. Ideally, each link would be measured at exactly the same time. However, this is technically not feasible as packets would collide and not be received. Instead, each module sends periodically every 25 ms (40 Hz) leading to an equidistant continuous sampling of the space covered by this 1-to-n broadcast link. Using this scheme increasing the number of senders is realized by intervening their periodic measurements. Given the previously defined time slot length we may use up to 10 transceivers before the first transceiver sends again. In order to achieve a stable continuous sampling a fixed sending sequence is required. This sequence can then be used to handle channel access using time multiplexing. Figure 72 illustrates the TDMA scheme for interleaved periodic transmission for these measurements.

IEEE 802.15.4 offers guaranteed time slots (GTS) for timing transmissions. Unfortunately, this is only available for unicast communication with a coordinator device. But only a single coordinator can exist in an IEEE 802.15.4 network. For this reason we needed to implement our own time division multiple access (TDMA) algorithm which handles channel access. When analyzing the clock drift of the transceiver modules relative to a specific transceiver clock (defined as the global

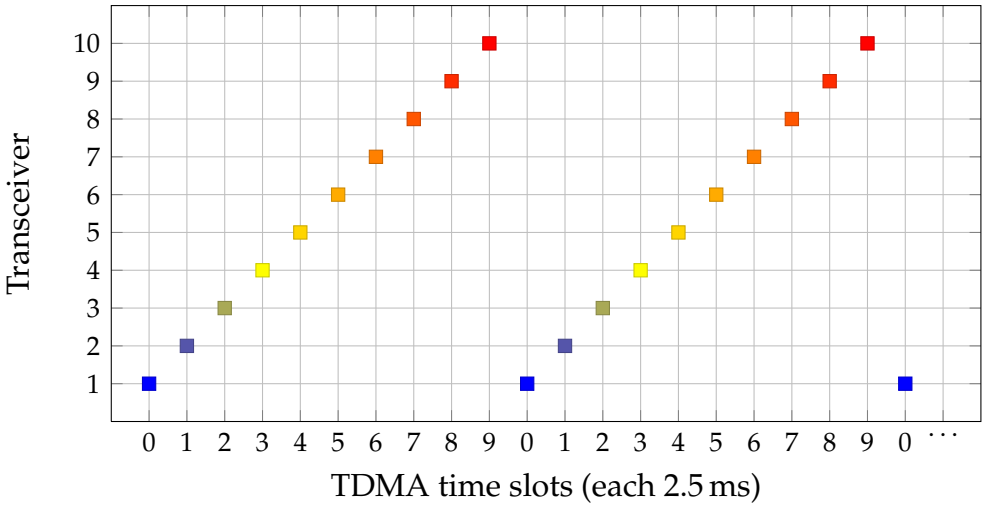



Figure 72.: Interleaved sending of packets for a Radio Sensor using 10 transceivers.

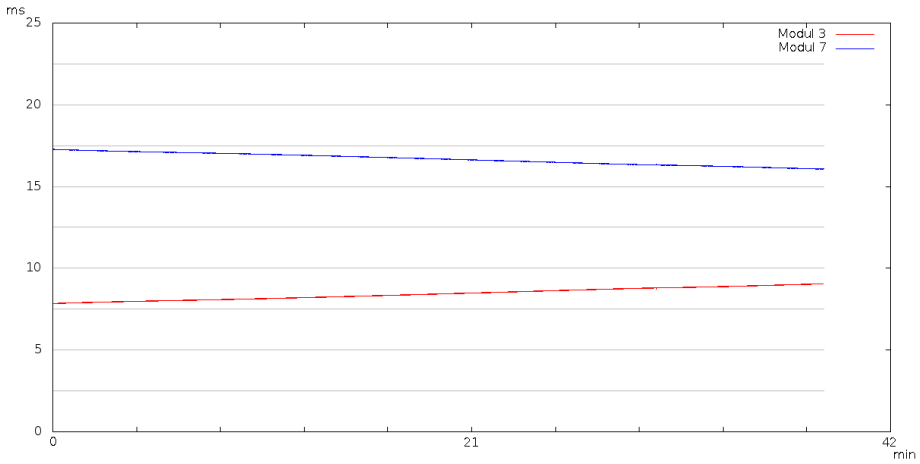



Figure 73.: Clock drift of two transceivers without synchronization.

time) we find that only after multiple minutes a module leaves its specified time slot. Figure 73 shows an illustration of the drifting behavior of two modules. Each grid line indicates the beginning of a 2.5 ms time slot of a module. We find after about 25 minutes the lower module has drifted in the middle between its own time slot at 7.5 ms and the next time slot at 10 ms.

Thus, clock drift becomes problematic only after multiple TDMA rounds. Still, we will attempt to correct clock drift in every round. This way, even if synchronization packets are not received for multiple successive rounds the system will continue working.

We implemented the synchronization by aligning the clocks of all transceiver nodes to the clock of a single master transceiver which defines the global clock. Alignment is conducted by comparing the master clock tick time for sending a broadcast packet to the clock tick time of each individual node for receiving the same packet.

As previously stated, the delay between send and receive is beneath the time resolution of the transceiver and can therefore be seen as instantaneous. Thus, the difference between master and node clocks is used to correct the local node tick count such that each node may send its next packets earlier or later depending if their own clock is either faster or slower than the master clock. While a clock tick on the master and a clock tick on a node has a different meaning (e.g. a master tick may be 1 $\mu s$ while a node tick may be 2 $\mu s$) this approach allows for synchronization as it is conducted continuously and hence the error does not sum up.

In addition, every time one of the nodes sends a packet each of the other nodes measures the received signal strength and puts it into the send buffer for its next packet. This way all nodes possess the complete measurement of the Radio Sensor. Thus, the Radio Sensor data may be analyzed by tapping into any of the nodes.

Given a working time synchronization (as shown in the next section). We can now sample the test bed area using 40 Hz for up to 10 transceivers.

*Extension for Frequency Diversity*

We extended the described sampling algorithm for frequency diversity by reducing the number of transceivers and re-using the free TDMA slots for sending broadcast packets on different frequency channels. As a frequency change happens in constant time and introduces no latency we can measure three frequencies with 40 Hz with four transceivers. Figure 74 visualizes the sequence.

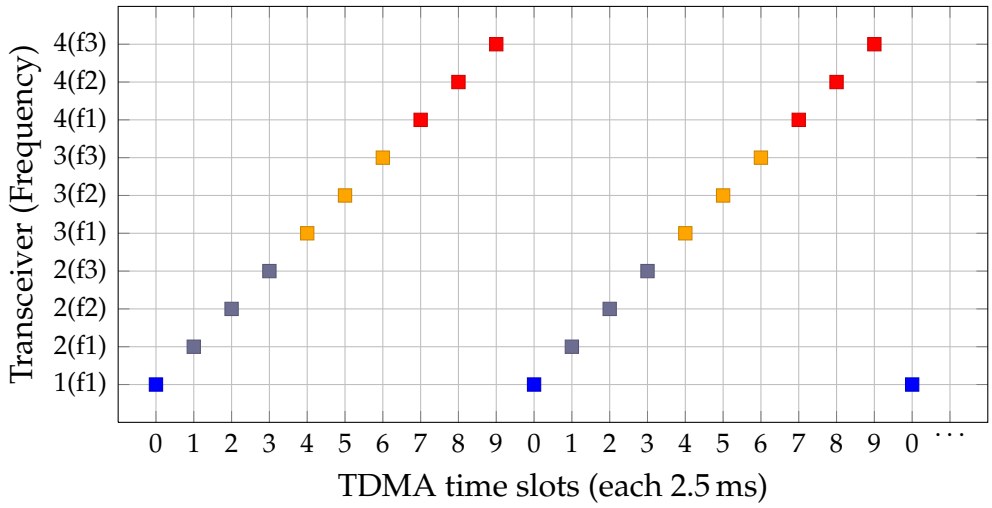



Figure 74.: Interleaved sending of packets for a Radio Sensor using 4 transceivers with frequency diversity (3 frequencies).

While the master (transceiver 1 in the figure) only sends once in one TDMA round, all other transceivers send thrice. After a transceiver (e.g. 2) sends in its original slot (e.g. 1) on frequency 1, it changes to frequency 2 for the following time slot (e.g. 2). After receiving the broadcast packet on frequency 1, all other transceivers also change to frequency 2. Then the transceiver sends another

broadcast on frequency 2 and changes to frequency 3. Again the other transceivers also change to frequency 3 after receiving the packet. The process is repeated a last time for frequency 3. After receiving this third packet all transceivers change back to frequency 1. The next transceiver (e.g. 3) is now triggered by its timer and starts the same process in its first time slot (e.g. 4). This way we can measure all links on three different carrier frequencies while keeping the required sample rate.

### A.4.4 *Validation*

In this section we present the validation of the synchronization algorithm by running the Radio Sensor for a longer period of time using the maximum number of transceivers. Figure 75 shows the results. In the figure each curve visualizes all

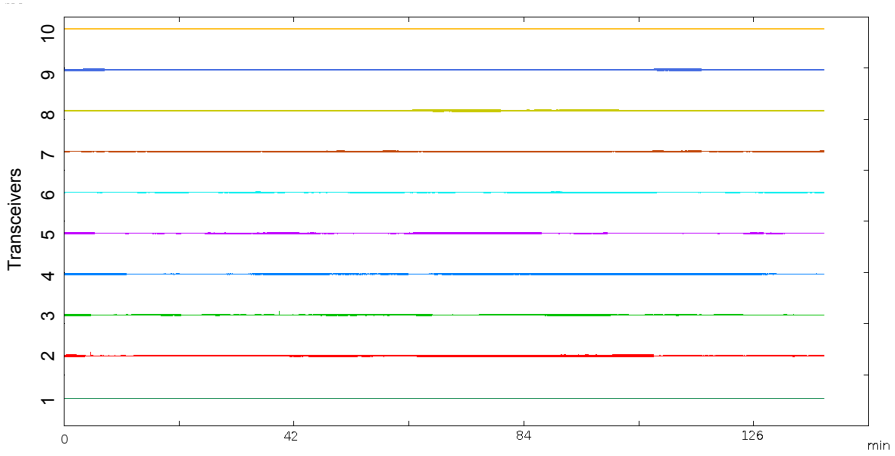



Figure 75.: Synchronized TDMA

points in time when a transceiver sent its broadcast packet relative to the master within one TDMA round (25 ms) over the course of more than 2 h. The curve at the bottom shows the master. We observed that while the exact send times differed slightly over the duration of the experiment these errors did not add up. Thus, assuming regular packet receptions by all nodes the Radio Sensor can run infinitely.

## A.5 DISCUSSION

### A.5.1 *Sampling Algorithm*

The presented sampling algorithm is similar to the algorithm presented by Kaltiokallio et al. [40] but uses shorter slot times, thereby enabling the incorporation of more transceiver modules and frequency diversity. However, slot time was reduced by switching off certain IEEE 802.15.4 features, therefore we cannot directly compare to the previous work. Additionally, the send delay is an aspect of the vendor hardware/firmware implementation. Thus, other hardware may enable a larger number of transceivers or frequencies with the same algorithm. We further

showed that clock drift does not effect sampling directly but only after a certain time. Hence, we further consider the implemented Radio Sensor algorithm robust as it does not break even when missing multiple packets in a row.

For instance, if a transceiver is covered by a subject standing in close proximity. However, in contrast to the SPIN algorithm by Wilson and Patwari [91] the system is not easily extensible to more nodes. This is trade off which comes with the guaranteed sample rate: additional nodes would not fit in the TDMA cycle.

### A.5.2 *Compatibility with the Smart Home*

The presented Radio Sensor was developed based on a number of requirements. These were defined in order to minimize uncertainty and conduct measurements influenced primarily by human physical activities. However, with these requirements and the presented implementation, we have departed from one of the visions of device-free sensing: the re-use of existing wireless infrastructure. For instance, our current implementation occupies the wireless channel completely with packet traffic, hardly allowing any other IEEE 802.15.4 communication (on the same channels). IEEE 802.15.4 was originally intended for limited bandwidth applications such as periodically transmitting temperature values. Thus, we can imagine that such information may be transferred on top of a Radio Sensor infrastructure e.g. by packing this data into a IEEE 802.15.4 packet along with the Radio Sensor measurements. However, the disabled retransmission in case of lost packets may be an issue. One workaround is to transmit a measurement multiple times in consecutive broadcast packets.

Another aspect is the continuous emission of electromagnetic radiation by the sensor. The current system is designed to sample fast enough for any human activity. But most activities of daily living only create frequencies of 0.3-3.5 Hz [13]. Hence, investigations may show that a Radio Sensor operating with reduced sampling frequency and emissions is sufficient. In addition, it should be considered that IEEE 802.15.4 is a low power technology. Its typically transmission power is on the order of 1 mW, which is the same for Bluetooth Low Energy. In contrast, WiFi typically operates in the range of 100 mW and the popular DECT phones even use 250 mW. Of course, it also depends how this energy is dissipated. But typically omnidirectional antennas are employed in all of these technologies. Thus, the actual energy received by an individual is only a fraction of this power.

## A.6 CONCLUSION

In this chapter, we have described the development of the Radio Sensor. The Radio Sensor is based on the IEEE 802.15.4 protocol, a common protocol in Smart Home communication. The sensor uses a highly synchronized TDMA allowing 40 Hz sampling between all radio transceivers which comprise the sensor. The current algorithm allows the sensor to either operate with four transceivers sending and receiving on three frequencies or 10 transceivers operating on a single frequency.

Therewith, it allows the double number of transceivers than a previously described algorithm which only operated on a single frequency. The algorithm performed well in a validation test and is suited for test bed installations for the experiments conducted in this thesis. Since it uses a custom communication scheme it is not "drop in" compatible to Smart Home infrastructures. However, it may be employed with modifications e.g. regarding the sampling frequency or by inclusion of Smart Home appliance traffic into Radio Sensor packets.

With the successful evaluation of the Radio Sensor we consider this requirement for a practical Activity Recognition system as described in Chapter 5 satisfied.

# BIBLIOGRAPHY


[1] H. Abdel-Nasser, R. Samir, I. Sabek, and M. Youssef. MonoPHY: Monostream-based device-free WLAN localization via physical layer information. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 4546–4551, April 2013.

[2] Heba Abdelnasser, Moustafa Youssef, and Khaled Harras, A. WiGest: A Ubiquitous WiFi-based Gesture Recognition System. In *INFOCOM, 2015 Proceedings IEEE*. IEEE, 2015.

[3] Shigeo Abe. *Support vector machines for pattern classification*, volume 53. Springer, 2005.

[4] William Ablondi. 2014 Smart Home Systems and Services Forecast: Germany, Mar 2014. Report.

[5] A. Al-Husseiny and M. Youssef. RF-Based Traffic Detection and Identification. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5, Sept 2012.

[6] Heba Aly and Moustafa Youssef. New Insights into Wifi-based Device-free Localization. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 541–548, New York, NY, USA, 2013. ACM.

[7] Moustafa Alzantot and Moustafa Youssef. CrowdInside: Automatic Construction of Indoor Floorplans. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pages 99–108, New York, NY, USA, 2012. ACM.

[8] Ling Bao and Stephen Intille. Activity Recognition from User-Annotated Acceleration Data. In Alois Ferscha and Friedemann Mattern, editors, *Pervasive Computing*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin / Heidelberg, 2004.

[9] Martin Berchtold, Matthias Budde, Dawud Gordon, Hedda Rahel Schmidtke, and Michael Beigl. Actiserv: Activity recognition service for mobile phones. In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–8. IEEE, 2010.

[10] Christopher M Bishop. *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[11] M. Bocca, O. Kaltiokallio, N. Patwari, and S. Venkatasubramanian. Multiple Target Tracking with RF Sensor Networks. *Mobile Computing, IEEE Transactions on*, 13(8):1787–1800, Aug 2014.

[12] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 1992.

[13] C.V.C. Bouten, K.T.M. Koekkoek, M. Verduin, R. Kodde, and J.D. Janssen. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*, 44(3):136 –147, March 1997.

[14] Raymond C. Browning, Emily A. Baker, Jessica A. Herron, and Rodger Kram. Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of Applied Physiology*, 100(2):390–398, 2006.

[15] Joseph Chamie. World population ageing. Technical report, Technical report, Department of Economic and Social Affairs, Population Division, United Nations, 2009.

[16] Liming Chen, Chris D Nugent, Jit Biswas, and Jesse Hoey. *Activity recognition in pervasive intelligent environments*, volume 4. Springer Science & Business Media, 2011.

[17] Jin Cheng and Thomas Kunz. A survey on smart home networking. *Carleton University, Systems and Computer Engineering, Technical Report SCE-09-10*, 2009.

[18] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*, 13(1):21–27, 1967.

[19] Brian P Crow, Indra Widjaja, Jeong Geun Kim, and Prescott T Sakai. IEEE 802.11 wireless local area networks. *Communications Magazine, IEEE*, 35(9):116–126, 1997.

[20] Yan Cui, Will Chang, Tobias Nöll, and Didier Stricker. KinectAvatar: Fully Automatic Body Capture Using a Single Kinect. In *ACCV Workshop on Color Depth Fusion in Computer vision*, 2012.

[21] Janez Demšar, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevar, Mitar Milutinovič, Martin Možina, Matija Polajnar, Marko Toplak, Anže Starič, Miha Štajdohar, Lan Umek, Lan Žagar, Jure Žbontar, Marinka Žitnik, and Blaž Zupan. Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14:2349–2353, 2013.

[22] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.

[23] Ludwig Fahrmeir, Rita Künstler, Iris Pigeot, and Gerhard Tutz. *Statistik*. Springer-Verlag, 2007.

[24] Shahin Farahani. *ZigBee wireless networks and transceivers*. Newnes, 2011.

[25] Jon E Froehlich, Eric Larson, Tim Campbell, Conor Haggerty, James Fogarty, and Shwetak N Patel. HydroSense: infrastructure-mediated single-point sensing of whole-home water activity. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 235–244. ACM, 2009.

[26] Dawud Gordon, Jan-Hendrik Hanne, Martin Berchtold, Takashi Miyaki, and Michael Beigl. An experiment in hierarchical recognition of group activities using wearable sensors. In *Modeling and Using Context*, pages 104–107. Springer, 2011.

[27] Dawud Gordon, Hedda Schmidtke, and Michael Beigl. Introducing New Sensors for Activity Recognition. In *In How To Do Good Research In Activity Recognition: Experimental methodology, performance evaluation and reproducibility*, Workshop in conjunction with Pervasive 2010, 2010.

[28] Sidhant Gupta, Matthew S Reynolds, and Shwetak N Patel. ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 139–148. ACM, 2010.

[29] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. Tool Release: Gathering 802.11N Traces with Channel State Information. *SIGCOMM Comput. Commun. Rev.*, 41(1):53–53, January 2011.

[30] Weihua He, Yongcai Guo, Chao Gao, and Xinke Li. Recognition of human activities with wearable sensors. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–13, 2012.

[31] Khaled Harras Heba Abdelnasir and and Moustafa Youssef. UbiBreathe: A Ubiquitous non-Invasive WiFi-based Breathing Estimator. In *ACM Mobihoc*. ACM, 2015.

[32] Rüdiger Hoffmann and Matthias Wolff. *Intelligente Signalverarbeitung 1: Signalanalyse*. Springer-Verlag, 2014.

[33] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification, 2003.

[34] Donny Huang, Rajalakshmi Nandakumar, and Shyamnath Gollakota. Feasibility and Limits of Wi-Fi Imaging. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, pages 266–279. ACM, 2014.

[35] IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs). *IEEE Std 802.15.4-2003*, pages 1–670, 2003.

[36] Digi International. XBee / XBee-PRO RF Modules Product Manual, Feb 2015. Revision S, Part number 900000982 S, Datasheet.

[37] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pages 559–568, New York, NY, USA, 2011. ACM.

[38] George H John, Ron Kohavi, Karl Pfleger, et al. Irrelevant features and the subset selection problem. In *Machine Learning: Proceedings of the Eleventh International Conference*, pages 121–129, 1994.

[39] Henry S. Kahn, Qiuping Gu, Kai McKeever Bullard, David S. Freedman, Namanjeet Ahluwalia, and Cynthia L. Ogden. Population Distribution of the Sagittal Abdominal Diameter (SAD) from a Representative Sample of US Adults: Comparison of SAD, Waist Circumference and Body Mass Index for Identifying Dysglycemia. *PLoS ONE*, 9(10):e108707, 10 2014.

[40] O. Kaltiokallio and M. Bocca. Real-Time Intrusion Detection and Tracking in Indoor Environment through Distributed RSSI Processing. In *Embedded and Real-Time Computing Systems and Applications (RTCSA), 2011 IEEE 17th International Conference on*, volume 1, pages 61–70, Aug 2011.

[41] O. Kaltiokallio, M. Bocca, and N. Patwari. Enhancing the accuracy of radio tomographic imaging using channel diversity. In *Mobile Adhoc and Sensor Systems (MASS), 2012 IEEE 9th International Conference on*, pages 254–262, Oct 2012.

[42] O. Kaltiokallio, M. Bocca, and N. Patwari. A Fade Level-Based Spatial Model for Radio Tomographic Imaging. *Mobile Computing, IEEE Transactions on*, 13(6):1159–1172, June 2014.

[43] N. Kassem, A.E. Kosba, and M. Youssef. RF-Based Vehicle Detection and Speed Estimation. In *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, pages 1–5, May 2012.

[44] Sidney Katz, Amasa B Ford, Roland W Moskowitz, Beverly A Jackson, and Marjorie W Jaffe. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *Jama*, 185(12):914–919, 1963.

[45] A.E. Kosba, A. Saeed, and M. Youssef. RASID: A robust WLAN device-free passive motion detection system. In *Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*, pages 180–189, March 2012.

[46] Ahmed E Kosba, Ahmed Abdelkader, and Moustafa Youssef. Analysis of a device-free passive tracking system in typical wireless environments. In *New Technologies, Mobility and Security (NTMS), 2009 3rd International Conference on*, pages 1–5. IEEE, 2009.

[47] Marc Kurz, Gerold Holzl, Alois Ferscha, Alberto Calatroni, Daniel Roggen, Gerhard Troster, Hesam Sagha, Ricardo Chavarriaga, Jose del R Millan, David Bannach, et al. The opportunity framework and data processing ecosystem for opportunistic activity and context recognition. *International Journal of Sensors Wireless Communications and Control*, 1(2):102–125, 2011.

[48] DavidD. Lewis. Naive (Bayes) at forty: The independence assumption in information retrieval. In Claire Nédellec and Céline Rouveirol, editors, *Machine Learning: ECML-98*, volume 1398 of *Lecture Notes in Computer Science*, pages 4–15. Springer Berlin Heidelberg, 1998.

[49] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. Wi-Sleep: Contactless Sleep Monitoring via WiFi Signals. In *Real-Time Systems Symposium (RTSS), 2014 IEEE*, pages 346–355, Dec 2014.

[50] Jennic Ltd. Data Sheet: JN5139-001 and JN5139-Z01, May 2009. Version 1.8, JN-DS-JN5139 1v8, Data Sheet.

[51] Brad Mager, Neal Patwari, and Maurizio Bocca. Fall detection using RF sensor networks. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, pages 3472–3476. IEEE, 2013.

[52] Prasant Misra, Nadeem Ahmed, Diethelm Ostry, and Sanjay Jha. Characterization of asymmetry in low-power wireless links: an empirical study. In *Proceedings of the 12th international conference on Distributed computing and networking*, ICDCN'11, pages 340–351, Berlin, Heidelberg, 2011. Springer-Verlag.

[53] May Moussa and Moustafa Youssef. Smart Devices for Smart Environments: Device-free Passive Detection in Real Environments. In *IEEE International Conference on Pervasive Computing and Communications*, pages 1–6, 2009.

[54] Michael Mozer. *Lessons from an adaptive house*. PhD thesis, Architectural Engineering, 2004.

[55] Bojan Mrazovac, Milan Z Bjelica, Dragan Kukolj, Sasa Vukosavljev, and Branislav M Todorovic. System design for passive human detection using principal components of the signal strength space. In *Engineering of Computer Based Systems (ECBS), 2012 IEEE 19th International Conference and Workshops on*, pages 164–172. IEEE, 2012.

[56] Elizabeth D Mynatt and Wendy A Rogers. Developing technology to support the functional independence of older adults. *Ageing International*, 27(1):24–41, 2001.

[57] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *Knowledge and Data Engineering, IEEE Transactions on*, 22(10):1345–1359, 2010.

[58] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca. Breathfinding: A Wireless Network That Monitors and Locates Breathing in a Home. *Selected Topics in Signal Processing, IEEE Journal of*, 8(1):30–42, Feb 2014.

[59] N. Patwari and J. Wilson. Spatial Models for Human Motion-Induced Signal Strength Variance on Static Links. *Information Forensics and Security, IEEE Transactions on*, 6(3):791 –802, sept. 2011.

[60] Neal Patwari, Joey Wilson, Sai Ananthanarayanan, Sneha K Kasera, and Dwayne R Westenskow. Monitoring breathing via signal strength in wireless networks. *Mobile Computing, IEEE Transactions on*, 13(8):1774–1786, 2014.

[61] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. Whole-home Gesture Recognition Using Wireless Signals. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*, MobiCom '13, pages 27–38, New York, NY, USA, 2013. ACM.

[62] J Ross Quinlan. *C4. 5: programs for machine learning*. Elsevier, 2014.

[63] Theodore S. Rappaport. *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall, 2 edition, January 2002.

[64] Markus Reschke, Johannes Starosta, Sebastian Schwarzl, and Stephan Sigg. Situation awareness based on channel measurements. In *Proceedings of the fourth Conference on Context Awareness for Proactive Systems (CAPS)*, 2011.

[65] Bernard Rosner. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics*, 25(2):165–172, 1983.

[66] I. Sabek, M. Youssef, and A.V. Vasilakos. ACE: An Accurate and Efficient Multi-Entity Device-Free WLAN Localization System. *Mobile Computing, IEEE Transactions on*, PP(99):1–1, 2014.

[67] A. Saeed, A.E. Kosba, and M. Youssef. Ichnaea: A Low-Overhead Robust WLAN Device-Free Passive Localization System. *Selected Topics in Signal Processing, IEEE Journal of*, 8(1):5–15, Feb 2014.

[68] Albrecht Schmidt. Implicit human computer interaction through context. *Personal technologies*, 4(2-3):191–199, 2000.

[69] Albrecht Schmidt, Michael Beigl, and Hans-Werner Gellersen. There is more to context than location. *Computers & Graphics*, 23:893–902, Dec 1999.

[70] Markus Scholz, Dawud Gordon, Leonardo Ramirez, Stephan Sigg, Tobias Dyrks, and Michael Beigl. A concept for support of firefighter frontline communication. *Future Internet*, 5(2):113–127, 2013.

[71] Markus Scholz, Lukas Kohout, Matthias Horne, Matthias Budde, Michael Beigl, and Moustafa Youssef. Device-Free Radio-based Low Overhead Identification of Subject Classes. In *Proceedings of the 2nd Workshop on Physical Analytics (WPA-15), co-located with ACM Mobisys*, volume 2015, 2015.

[72] Markus Scholz, Till Riedel, and Christian Decker. A flexible architecture for a robust indoor navigation support device for firefighters. In *Networked Sensing Systems (INSS), 2010 Seventh International Conference on*, pages 227–232. IEEE, 2010.

[73] Markus Scholz, Till Riedel, Mario Hock, and Michael Beigl. Device-free and Device-bound Activity Recognition Using Radio Signal Strength. In *Proceedings of the 4th Augmented Human International Conference*, AH '13, pages 100–107, New York, NY, USA, 2013. ACM.

[74] Markus Scholz, Stefan Sigg, Gerrit Bagschik, Toni Guenther, Georg von Zengen, Dimana Shishkova, Yusheng Ji, and Michael Beigl. SenseWaves: Radiowaves for context recognition, Video submission. In *Pervasive*, 2011.

[75] Markus Scholz, Stephan Sigg, Hedda R Schmidtke, and Michael Beigl. Challenges for device-free radio-based activity recognition. In *Proceedings of the 3rd workshop on Context Systems, Design, Evaluation and Optimisation (CoSDEO 2011), in Conjunction with MobiQuitous*, volume 2011, 2011.

[76] Erhard Schubert and Markus Scholz. Evaluation of wireless sensor technologies in a firefighting environment. In *Networked Sensing Systems (INSS), 2010 Seventh International Conference on*, pages 157–160. IEEE, 2010.

[77] Shuyu Shi, Stephan Sigg, and Yusheng Ji. Passive Detection of Situations from Ambient FM-radio Signals. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 1049–1053, New York, NY, USA, 2012. ACM.

[78] Shuyu Shi, Stephan Sigg, and Yusheng Ji. ActiviTune: A Multi-stage System for Activity Recognition of Passive Entities from Ambient FM-Radio Signals. In Kui Ren, Xue Liu, Weifa Liang, Ming Xu, Xiaohua Jia, and Kai Xing, editors, *Wireless Algorithms, Systems, and Applications*, volume 7992 of *Lecture Notes in Computer Science*, pages 221–232. Springer Berlin Heidelberg, 2013.

[79] Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Commun. ACM*, 56(1):116–124, January 2013.

[80] Bluetooth SIG. Specification of the Bluetooth System Version 4.2, Dec 2014.

[81] S. Sigg, U. Blanke, and G. Troster. The telepathic phone: Frictionless activity recognition from WiFi-RSSI. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, pages 148–155, March 2014.

[82] S. Sigg, M. Scholz, Shuyu Shi, Yusheng Ji, and M. Beigl. RF-Sensing of Activities from Non-Cooperative Subjects in Device-Free Recognition Systems Using Ambient and Local Signals. *Mobile Computing, IEEE Transactions on*, 13(4):907–920, April 2014.

[83] Stephan Sigg, Mario Hock, Markus Scholz, Gerhard Tröster, Lars Wolf, Yusheng Ji, and Michael Beigl. Passive, Device-Free Recognition on Your Mobile Phone: Tools, Features and a Case Study. In Ivan Stojmenovic, Zixue Cheng, and Song Guo, editors, *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, volume 131 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 435–446. Springer International Publishing, 2014.

[84] Stephan Sigg, Shuyu Shi, and Yusheng Ji. Rf-based device-free recognition of simultaneously conducted activities. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 531–540. ACM, 2013.

[85] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, 2009.

[86] Fawcett T. ROC graphs: Notes and practical considerations for data mining researchers. Technical Report HPL-2003-4, HP Laboratories, Palo Alto, CA, USA, Jan 2003.

[87] Texas Instruments Incorporated. A USB-Enabled System-On-Chip Solution for 2.4-GHz IEEE 802.15.4 and ZigBee Applications, Jun 2010. SWRS086A, Revised June 2010, Data Sheet.

[88] Atmel Corporation USA. Low Power, 2.4GHz Transceiver for ZigBee, RF4CE, IEEE 802.15.4, 6LoWPAN, and ISM Applications, Jul 2014. Atmel 8351E MCU Wireless AT86RF233 Datasheet.

[89] Guanhua Wang, Yongpan Zou, Zimu Zhou, Kaishun Wu, and Lionel M. Ni. We Can Hear You with Wi-Fi! In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom '14, pages 593–604, New York, NY, USA, 2014. ACM.

[90] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. E-eyes: Device-free Location-oriented Activity Identification Using Fine-grained WiFi Signatures. In *Proceedings of the 20th Annual International*

*Conference on Mobile Computing and Networking*, MobiCom '14, pages 617–628, New York, NY, USA, 2014. ACM.

[91] Anthony Joseph Wilson and Neal K Patwari. Systems and methods of device-free motion detection and presence sensing, March 14 2014. US Patent App. 14/212,447.

[92] Joey Wilson and Neal Patwari. Radio Tomographic Imaging with Wireless Networks. *IEEE Transactions on Mobile Computing*, 9:621–632, 2010.

[93] K. Woyach, D. Puccinelli, and M. Haenggi. Sensorless sensing in wireless networks: implementation and measurements. In *Proceedings of the Second International Workshop on Wireless Network Measurement (WiNMee)*, 2006.

[94] Jiang Xiao, Kaishun Wu, Youwen Yi, Lu Wang, and Lionel M. Ni. FIMD: Fine-grained Device-free Motion Detection. In *Proceedings of the 2012 IEEE 18th International Conference on Parallel and Distributed Systems*, ICPADS '12, pages 229–235, Washington, DC, USA, 2012. IEEE Computer Society.

[95] Chenren Xu, Bernhard Firner, Robert S. Moore, Yanyong Zhang, Wade Trappe, Richard Howard, Feixiong Zhang, and Ning An. SCPL: Indoor Device-free Multi-subject Counting and Localization Using Radio Signal Strength. In *Proceedings of the 12th International Conference on Information Processing in Sensor Networks*, IPSN '13, pages 79–90, New York, NY, USA, 2013. ACM.

[96] Yaser Yacoob and Michael J Black. Parameterized modeling and recognition of activities. In *Computer Vision, 1998. Sixth International Conference on*, pages 120–127. IEEE, 1998.

[97] Jie Yang, Yong Ge, Hui Xiong, Yingying Chen, and Hongbo Liu. Performing joint learning for passive intrusion detection in pervasive wireless environments. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[98] Qingming Yao, Hui Gao, Bin Liu, and Fei-Yue Wang. MODEL: moving object detection and localization in wireless networks based on small-scale fading. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pages 451–452. ACM, 2008.

[99] Moustafa Youssef, Matthew Mah, and Ashok Agrawala. Challenges: device-free passive localization for wireless environments. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, MobiCom '07, pages 222–229, New York, NY, USA, 2007. ACM.

[100] Dian Zhang, Kezhong Lu, Rui Mao, Yuhong Feng, Yunhuai Liu, Zhong Ming, and L.M. Ni. Fine-Grained Localization for Multiple Transceiver-Free Objects by using RF-Based Technologies. *Parallel and Distributed Systems, IEEE Transactions on*, 25(6):1464–1475, June 2014.

[101] Dian Zhang, Jian Ma, Quanbin Chen, and Lionel M. Ni. An RF-Based System for Tracking Transceiver-Free Objects. In *Pervasive Computing and Communications, 2007. PerCom '07. Fifth Annual IEEE International Conference on*, pages 135 –144, march 2007.

[102] Dian Zhang and L.M. Ni. Dynamic clustering for tracking multiple transceiver-free objects. In *IEEE International Conference on Pervasive Computing and Communications (PerCom 2009)*, march 2009.

[103] Zimu Zhou, Zheng Yang, Chenshu Wu, Longfei Shangguan, and Yunhao Liu. Towards omnidirectional passive human detection. In *INFOCOM, 2013 Proceedings IEEE*, pages 3057–3065. IEEE, 2013.

# Acknowledgements

Firstly, I would like to thank my primary advisor Professor Michael Beigl for the possibility to write this thesis at his chair. I am very grateful for his constructive comments and his support and advice over all my years as a PhD student at his institute.

I further would like to thank my secondary advisor, Professor Moustafa A. Youssef, for agreeing to co-advise this thesis. Professor Youssef is one of the pioneers of the device-free, radio-based research field. It was an honour and a great pleasure to work with him. I especially enjoyed the insightful, amicable and clement discussions when he visited TECO/KIT in 2014. I have learned a great lot during this time.

I also would like to thank Stephan Sigg with whom I worked together on several research papers and who introduced me to this exciting research field.

I further would like to thank my colleagues and friends at TECO. They make this an outstanding environment for creativity and scientific work. I especially like to mention Till Riedel, Antonios Karatzoglou, Matthias Berning, Matthias Budde and Dawud Gordon. A special thank you also goes to TECO's very kind and most helpful secretary Irina Schierholz for supporting me with the organizational aspects of the thesis.

Writing a dissertation is hardly possible without the hard work of many Bachelor, Master and Diploma students. Thank you to all of you who have helped me along with various implementations and experiments.

Furthermore, I would like to thank my parents and my siblings for their endless support and advise throughout all the years, even when I was about to loose faith in this venture.

Finally, I would like to thank Kathleen, my dear partner, for her advice and support, for helping me recharge when I felt drained and for proofreading the complete manuscript for the final correction iteration.