Mining Contexts of Passive Entities from Radio Frequency Sources

Shuyu Shi

Doctor of Philosophy

Department of Informatics
School of Multidisciplinary Sciences
SOKENDAI (The Graduate University for Advanced Studies)

Mining Contexts of Passive Entities from Radio Frequency Sources

Shuyu Shi
Department of Informatics
SOKENDAI(The Graduate University for Advanced Studies)

This dissertation is submitted for the degree of
Doctor of Philosophy

2016

I would like to dedicate this thesis to my loving parents ...

Acknowledgements

During the past years of my graduate studies, there are so many people I would like to thank for their help. First and foremost, I am immensely grateful to Professor Yusheng Ji, my amazing and inspiring advisor. She have guided me through the research process, taught me how to frame problems and analyze research, and have supported me no matter what over the past five years. Beyond this, her unfailingly positive attitude and her excellent advice renewed my spirits when I was low on energy; I always walked out of our meetings with a newfound vigor and enthusiasm for my work. And her ability to get to the heart of an issue and distill the key points has helped me focus my work. These are only a few ways in which I have benefited tremendously from both relationships. My Ph.D journey won't be so enlightened, enjoyful and fruitful without her.

I would also like to thank my subadvisor Stephan Sigg for his support and time he spent to guide and teach me how to accomplish academic goal. I gained enormous benefits from his knowledge, ranging from hardware, wireless networking to academic writing. He always encourages me to try my best to understand the fundamentals when encountering any problem. Furthermore, his endless passion for academia always excites me.

I have enjoyed collaborating with many other professors, students, and researchers over the years. In particular, Wenjun Hu and Neal Patwari were great mentors. I was fortunate to work with Lin Chen, Kaigui Bian, Rui Li, Yu Mao, Jinyun Feng, Ruijian An, Hao Zhou and Wei Zhao over the years, and I wish we could have worked together more.

Most importantly, I'd like to thank my family. They have been there for me through every

stage of my life, and their encouragement and support have made all the difference.

Abstract

The vitality of detecting human context information lies in that it relates closely to almost all our daily aspects, such as locating where a person is, knowing what he/she is doing and identifying people nearby. In the past decade, research of context awareness has focused more deeply on studies of location awareness and activity recognition, since the two aspects are involved in a large number of practical applications, ranging from location-based services, medical care to social science. The recent decade, with the ubiquitous distribution of mobile device integrated with a suite of sensors including accelerometer, microphone, GPS, WiFi, camera, digital compass and gyroscope, the device can provide a variety of contexts, such as user's physical location, activities and orientation, etc. However, despite the huge progress in this research area, there is still much more to be investigated. For instance, 1) most of existing solutions require wearing a device, then one could ask what can be done with the absence of such a device; and 2) the accuracy of RF-based indoor localization is constrained by the unpredictable characteristics of radio signals in indoor environments and hence needs to be improved.

In this dissertation, we design two novel context-aware systems that can address the problems like those described in the previous section and learn the useful human context information such as activities, locations attention and velocity. More specific, we first present a passive FM radio source based activity recognition system. Then, we propose an WiFi based indoor device-free positioning and tracking scheme using fine-grained Channel State Information (CSI) as source information.

(1) ActiviTune: A Multi-stage framework for activity recognition from FM-radio signals

In activity recognition, the majority of frequently applied sensing technologies require prior installation and repeated calibration within a changing environment. Examples are sensors for motion detection such as acceleration-type sensors as well as video or RFID-based systems. Clearly, the installation effort might diminish by proper integration of sensors in everyday objects, such as clothing. This has the potential to seamlessly foster the distribution of sensing equipment but also implicates the potential for inaccurate placement, thus increasing the importance of frequent sensor calibration. An approach that mitigates the necessity to equip users and also the frequent recalibration of the system is to establish context awareness using to establish environmental modalities.

The RF-transceiver of electronic equipment might constitute such kind of sensor. Nearly all contemporary electronic devices contain some kind of interface to the RF-channel. We propose to use the RF-channel as a sensing source to detect activities of individuals by monitoring peculiarities of RF-based features induced by these very activities. Comparing with other RF-based approaches, ActiviTune has the advantage of neither installing a transmitter generating the signal nor equipping the monitored entities with any active component of the system. In this thesis, we test our system to distinct several static and dynamic activities, by which, an overall accuracy of over 80% can be attained. Also, we further conduct another case study to explore the feasibility of discriminating both activities and locations of a subject simultaneously.

Given the overabundance of information, attention management is of great interest to the research community of ubiquitous computing and context awareness. We believe our above proposed FM-based activity recognition framework can provide a general solution to a plethora of context-aware computing systems. In this thesis, we exemplify an application using ambient FM-radio signals in monitoring of attention. We analyze the features extracted

from ambient FM-radio signals to understand the locomotion of an individual passing by a poster and then take it as a hint to imply the level of attention attracted by that poster.

(2) Mining Location Contexts Using CSI Information

Besides user's activity, context-aware systems are also concerned with learning the location of an entity, since it is a crucial type of context. In indoor environment, GPS-based techniques are not operational indoors. Accurate and ubiquitously available indoor localization is still an unsolved issue.

As the presence of a device-free subject can interfere with the ambient RF signals derived from radio sources, researchers have explored passive device-free (DfP) indoor localization, utilizing the change of RF signals induced by a monitored individual to estimate the person's location. However, due to the multipath effect and the limited granularity of Received Signal Strength Indicator (RSSI), there is still much room to improve the accuracy of indoor localization.

In this thesis, we propose a probabilistic fingerprinting-based technique for DfP indoor localization. Our system adopts CSI readings derived from off-the-shelf WiFi 802.11n wireless cards which can provide fine-grained subchannel measurements in the context of MIMO-OFDM PHY layer parameters. This complex channel information enables accurate localization of non-equipped individuals. Our scheme further boosts the localization efficiency by using principal component analysis (PCA) to identify the most relevant feature vectors. The experimental results demonstrate that our system can achieve an accuracy of over 92% and an error distance smaller than 0.5m. We also investigate the effect of other parameters on the performance of our system, including packet transmission rate, the number of links as well as the number of principle components.

Contents

Contents xi

List of Figures xv

List of Tables xvii

Nomenclature xviii

1 Introduction 1

1.1 Background and Motivation 1

1.2 Proposed Approaches 4

1.2.1 Mining Location and Activity Contexts from FM sources 5

1.2.1.1 ActiviTune: A Multi-stage framework for activity recognition from FM-radio signals 5

1.2.1.2 Use Case: an Attention Monitoring System from FM signals 5

1.2.2 Mining Location Contexts Using CSI Information 6

1.3 Contributions 7

1.4 Organization of this Dissertation 8

2 Related Work 9

2.1 Overview 9

2.2 Indoor Localization Using RF 10

2.2.1 Device-bound Localization 10
2.2.2 Device-free Localization 11
2.2.3 Device-free Tracking 12
2.3 Activity Recognition Using RF 12

3 Mining Location and Activity Contexts from FM Sources 17
3.1 Introduction 17
3.2 Background and Feasibility Study 19
3.2.1 Principles of RF signal propagation 20
3.2.2 Discussion of various RF-technologies for passive DFAR 21
3.2.3 Environment induced characteristics of RF signals 22
3.2.4 An application scenario for DFAR 24
3.3 FM-based Passive Activity Recognition Framework 24
3.3.1 A passive DFAR system 25
3.3.2 Features for passive DFAR 26
3.4 Experimental Evaluation 27
3.4.1 Recognition of activities from FM signals 28
3.4.2 Influence of distance 31
3.4.3 Alternative environments 32
3.4.4 Classification accuracy with accelerometer devices 35
3.4.5 Classification accuracy of active WSN based DFAR 36
3.4.6 Joint localization and activity recognition 39
3.5 A Use Case: An Attention Monitoring System from FM Signal 41
3.5.1 Overview 41
3.5.2 System design 42
3.5.3 Evaluation 43
3.5.3.1 State of the corridor 44

3.5.3.2 Focused attention towards specific frames 45
3.5.3.3 Tracking individuals in motion 46
3.5.3.4 Changes in the walking speed 47
3.5.3.5 Altering the count of receive devices 48
3.6 Summary 50

4 Mining Location Contexts Using CSI information 51

4.1 Introduction 51
4.2 Background and Motivation 53
4.2.1 Channel State Information 53
4.2.2 Unpredictable Nature of CSI in Indoor Environments 54
4.3 Localization System Overview 56
4.3.1 Problem Formulation 56
4.3.2 Probabilistic Method for Location Discrimination 58
4.3.3 Dimensionality Reduction 60
4.4 Experimental Setup 61
4.4.1 Hardware Description and Data Aggregation 61
4.4.2 Experiment Layouts 62
4.4.3 Performance Metrics 62
4.4.3.1 Cell Estimation Accuracy 62
4.4.3.2 Median Distance Error 62
4.4.3.3 Average Processing Time 64
4.4.4 Experimental Results 64
4.4.4.1 Comparing Various Localization Methods 64
4.4.4.2 Impact of Principal Components 65
4.4.4.3 Influence of Data Aggregation Parameters 66
4.4.4.4 Packet Reception Rate 66

4.4.4.5 Number of Links 67
4.4.4.6 Time of Collecting Data 68
4.5 Tracking With Bayesian Filtering 69
4.5.1 Kalman Filter Based Tracking System 69
4.5.1.1 Modeling 70
4.5.1.2 Update 74
4.5.2 Particle Filter Based Tracking System 75
4.6 tracking results 77
4.6.1 Data Set 77
4.6.2 Tracking Accuracy 78
4.7 Summary 79

5 Conclusion and Future Work **81**
5.1 Discussion 81
5.2 Conclusion 82
5.3 Future Work 83

References **85**

A Publication List **101**
A.1 Journal Publications 101
A.2 Conference and Workshop Publications 101
A.3 Joint Publications 102

List of Figures

3.1 Evolution of signal strength for all activities considered, performed by a single subject 23

3.2 An overview of the signal acquisition, processing and activity recognition of the passive DFAR system 25

3.3 Schematic illustration of the seminar room used in our case study 28

3.4 Features over 128 samples for 5 activities utilizing the raw signal strengths shown in fig 3.1 31

3.5 Activity recognition accuracy conducted by 3 subjects at area L_1 to L_6 by k-Nearest-Neighbor classifier 34

3.6 Visualization of object impact on the signals with different distance to the transceiver 35

3.7 Schematic illustration of the seminar room used in the case study 35

3.8 Sketch of the evaluation setting. The attention-monitoring system extracted features combining the data acquired by both USRP devices. 43

3.9 Comparison of the classification accuracy for the 4 cases described in section 3.5.3.1, section 3.5.3.2, section 3.5.3.3, and section 3.5.3.4 when only one receive device is utilized 49

4.1 Experimental indoor environment 55

4.2 the probability from all the 30 subchannels that the CSI measurement changes result in constructive or destructive effect respectively when a subject blocks the LOS path in the area of interest 56
4.3 Amplitude change of CSI readings in different subchannels 57
4.4 Layout sketches of our experimental indoor spaces. Environments (a), (b) and (c) are in the same building with walls mainly made of concrete. Environment (c) features a different penetration of walls which are made of wood and gypsum. 63
4.5 Comparison of cell estimation accuracy achieved by different indoor localization systems in four indoor environments 66
4.6 Comparison of median distance error achieved by different indoor localization systems in four indoor environments 67
4.7 Cell estimation accuracy with different proportion of principle components (PCs) in four indoor environments 68
4.8 Median distance error with different proportion of principle components (PCs) in four indoor environments 69
4.9 Average processing time to classify a test location with different proportion of principle components (PCs) in four indoor environments 70
4.10 Influence of packet reception rate on the performance of our indoor localization system 71
4.11 Influence of number of links on the performance of our indoor localization system 72
4.12 The cumulative distribution functions (CDFs) of tracking errors of our proposed tracking approaches in different indoor environments 80

List of Tables

3.1 Activities performed with respect to locations shown in figure 3.3. 29

3.2 Mean accuracy for one-stage method of the activities 'lying', 'standing', 'crawling', 'walking' and 'empty' from amplitude based features extracted from an FM-radio signal in the environment depicted in figure 3.3 30

3.3 Mean accuracy for two-stage method of the activities 'lying', 'standing', 'crawling', 'walking' and 'empty' from amplitude based features extracted from an FM-radio signal in the environment depicted in figure 3.3 32

3.4 Activities performed with respect to locations shown in figure 3.3. 32

3.5 Mean accuracy for two-stage method of the activities 'lying', 'standing', 'crawling', 'walking' and 'empty' from amplitude based features extracted from an FM-radio signal in the environment depicted in figure 3.3 33

3.6 activity recognition accuracies in comparison between RF sensor and accelerometer sensor performed in a conference room illustrated as fig 3.7 37

3.7 Class accuracies (true positive rates) achieved by a k-NN classification algorithm for an active WSN-based DFAR system when increasing the number of nodes in the experimentation area 38

3.8 Class accuracies (true positive rates) achieved by a DT classification algorithm for an active WSN-based DFAR system when increasing the number of nodes in the experimentation area 38

3.9 activity recognition accuracies in comparison between RF sensor and accelerometer sensor performed in a conference room illustrated as fig 3.7 39
3.10 Accuracy of joint localization and activity recognition for the activities 'lying', 'standing', 'walking', 'walking' and positions S_1 and S_2 in a corridor 40
3.11 Mean accuracy for the distinction of the corridor states 'empty', 'person standing' and 'person walking' 45
3.12 Mean accuracy for the distinction in front of which poster a person is standing 46
3.13 Mean accuracy for the distinction of walking at location A, B, C or D in the environment depicted in figure 3.8 47
3.14 Confusion matrices for the discrimination between walking speeds (0.5 m/s, 1 m/s, 2 m/s) achieved by k-NN and Decision Tree classifiers 48
3.15 Confusion matrices for the discrimination between walking speeds (0.5 m/s, 2 m/s) achieved by k-NN and Decision Tree classifiers 48

4.1 Default data aggregation parameters used in our experiments 73
4.2 Influence of time frame of data aggregation on the performance of our indoor localization system during training phase 73
4.3 Influence of time frame of data aggregation on the performance of our indoor localization system during test phase 74
4.4 Experimental parameters used in our tracking system 78
4.5 Average MAEs (in meters) from our proposed tracking approaches 79

Chapter 1

Introduction

This chapter provides some background to context-aware computing needed to understand my thesis and discusses the motivations that I consider to leverage ambient frequency radio (RF) signals for mining the location and activity context of an individual. Then, we briefly introduce our proposed solutions and contributions in the thesis. Finally, we present the organization of the rest of this dissertation.

1.1 Background and Motivation

The concept of *ubiquitous computing* or *ubicomp* is first introduced by Mark Weiser in his 1991 *Scientific American* article [88]. The vision of ubiquitous computing is to have mobile and embedded devices can communicate with each other anytime and everywhere in an environment, enabling computing technologies to support daily activities. According to the definition in [38], context-aware computing is one of distinct areas where the research of ubiquitous computing is focused.

The term context-aware computing was first coined by Schilit and Theimer [57], envisioning the use of context to allow ubiquitous computing applications to deliver the right services at the right time. In order to implement various applications of context-aware computing, we

must understand what context is and how it can be used. Schilit and Theimer [57] refer to "context" as location, identities of nearby people and objects, and changes to those objects. However, with this definition, it is difficult to determine whether a type of information not listed in the definition is context or not. Due to this limitation, in the work of [17], the context is defined as "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves." This definition can make it easier for an application to enumerate the context. For instance, for the application of an indoor mobile tour guide, the entities are the user, the application and the tour sites. Regarding the information of the presence of other people, since it can affect the the user's situation, it is treated as context.

Though as defined by Dey [17], the context can be regarded as "any information that can be used to characterize the situation of an entity", a majority of context-aware applications can be categorized into location awareness systems and activity recognition systems, since location and activity are of vital importance for an entity. Therefore, in this thesis, research and developments of context-aware computing are discussed with a special focus on mining the location and activity information of an entity.

Thanks to the proliferation of mobile and wearable devices and the pervasive presence of Internet and various wireless networks, these advances are leading the vision of ubiquitous computing into reality, that is computing and communication is not limited to high-end computers, but can also seamlessly extend to embedded and smart devices. A typical example is the smartphones integrated with various sensors including GPS, WiFi, camera, digital compass, gyroscope, accelerometer and microphone, by which we can acquire a large range of context-aware information, such as the walking velocity and step counting of an user using accelerometer, user's location in outdoor environment via GPS and heart rate monitoring of a subject. The omnipresence of mobile devices and rapid development in computing technolo-

gies have led to a plethora of new applications in recent years ranging from social networks, mobile health caring, intelligent transportation.

Still, several shortcomings of existing technology impede the development of localization and activity recognition systems capable of providing useful context information for users. In the design of such systems, the foremost issue is most applications assume that sensors which can aggregate information sources are attached to the sensed entities. For instance, step counting Apps are only reliable once the user is carrying the mobile device. These systems are defined as *active* approaches. However, in the study [49], Patel et al. report that users only have a share of 58 percentage of time when the mobile devices are within their arm reach. Regarding this challenge, new technology should be capable of sensing subjects without equipping them. Another issue to be addressed is the system cost and installation effort that can scale to a large area of interest. An example is the camera surveillance system, which can only monitor human activities or position of a subject in the line-of-sight. Obviously, the installation cost impedes the deployment at scale. An alternative should be to exploit the ubiquitously available sensing modalities to mitigate the installation effort. Furthermore, unlike outdoor localization via Global Positioning System (GPS), to acquire the context of an user's location in indoor environment, we should deal with the challenge of multipath propagation of radio signals. Given the aforementioned issues, there has recently been an increasing interest in research on unnoticeably learning contexts of entities using readily environmental sensing modalities including inferred, ultrasound, camera, sound, light and radio frequency (RF), which are defined as *passive* methods. Among these modalities, RF is most promising due to its ubiquity.

RF-based passive context-aware technologies leverage physical phenomena of RF propagation, i.e. diffraction, reflection and scattering, for learning the location or activity information of a (or several) subject(s). The first experiment in this research domain dates back to the work [93], demonstrating that the characteristics of RF signals change when the presence of

object interferes with these signal.

Based on the observation, a plethora of solutions have been proposed focusing on passive indoor localization from ambient RF sources. The received signal strength indicator (RSSI) based schemes are presented in the studies [43, 92, 101]. Since the RF devices of these systems operate at 2.4 GHz, other researchers [53] exploit FM radio for passive indoor localization owing to its lower frequency and hence better robustness to penetration, multipath and distance of transmission. More recently, passive localization systems based on fine-grained channel state information (CSI) have been designed for passive positioning of an unequipped subject [85, 95, 96]. Furthermore, human activity recognition and body motion detection schemes have been proposed in recent research projects leveraging ambient RF signals [3, 32, 50, 54, 59]. However, to the best of our knowledge, there is no system which detects the activity information of an entity from ambient FM radio sources. Regarding the research on passively mining the location information of a subject, despite the recent progress, there is still much room to improve the precision in this area. To this end, in this thesis, we firstly design a FM-radio based passive activity recognition framework. Furthermore, we propose a probabilistic fingerprint-based positioning solution using CSI readings.

1.2 Proposed Approaches

In this dissertation, we design two novel context-aware systems that can address the problems like those described in the previous section and learn the useful human context information such as activities, locations, attention and velocity. More specific, we first present a passive FM radio source based activity recognition system. Then, we propose a WiFi based indoor device-free positioning and tracking scheme using fine-grained Channel State Information (CSI) as source information.

1.2.1 Mining Location and Activity Contexts from FM sources

1.2.1.1 ActiviTune: A Multi-stage framework for activity recognition from FM-radio signals

In activity recognition, the majority of frequently applied sensing technologies require prior installation and repeated calibration within a changing environment. Examples are sensors for motion detection such as acceleration-type sensors [6, 40] as well as video or RFID-based systems [18]. Clearly, the installation effort might diminish by proper integration of sensors in everyday objects, such as clothing. This has the potential to seamlessly foster the distribution of sensing equipment but also implicates the potential for inaccurate placement, thus increasing the importance of frequent sensor calibration. An approach that mitigates the necessity to equip users and also the frequent recalibration of the system is to establish context awareness using to establish environmental modalities.

The RF-transceiver of electronic equipment might constitute such kind of sensor. Nearly all contemporary electronic devices contain some kind of interface to the RF-channel. We propose to use the RF-channel as a sensing source to detect activities of individuals by monitoring peculiarities of RF-based features induced by these very activities. Comparing with other RF-based approaches, ActiviTune has the advantage of neither installing a transmitter generating the signal nor equipping the monitored entities with any active component of the system. In this thesis, we test our system to distinct several static and dynamic activities, by which, an overall accuracy of over 80% can be attained. Also, we further conduct another case study to explore the feasibility of discriminating both activities and locations of a subject simultaneously.

1.2.1.2 Use Case: an Attention Monitoring System from FM signals

Nowadays, we are approaching the era of big data, since individuals can acquire floods of information originated from a variety of sources ranging from web-based news, social network

systems, mail systems to Internet Protocol TV (IPTV). Given the overabundance of information, attention management is of great interest to the research community of ubiquitous computing and context awareness.

We believe our above proposed FM-based activity recognition framework can provide a general solution to a plethora of context-aware computing systems. In this thesis, we exemplify an application using ambient FM-radio signals in monitoring of attention. We analyze the features extracted from ambient FM-radio signals to understand the locomotion of an individual passing by a poster and then take it as a hint to imply the level of attention attracted by that poster.

1.2.2 Mining Location Contexts Using CSI Information

Besides user's activity, context-aware systems are also concerned with learning the location of an entity, since it is a crucial type of context. In indoor environment, GPS-based techniques are not operational indoors. Accurate and ubiquitously available indoor localization is still an unsolved issue.

As the presence of a device-free subject can interfere with the ambient RF signals derived from radio sources, researchers have explored passive device-free (DfP) indoor localization [12, 46, 52], utilizing the change of RF signals induced by a monitored individual to estimate the person's location. However, due to the mulitipath effect and the limited granularity of Received Signal Strength Indicator (RSSI), there is still much room to improve the accuracy of indoor localization.

In this thesis, we propose a probabilistic fingerprinting-based technique for DfP indoor localization. Our system adopts CSI readings derived from off-the-shelf WiFi 802.11n wireless cards which can provide fine-grained subchannel measurements in the context of MIMO-OFDM PHY layer parameters. This complex channel information enables accurate localization of non-equipped individuals. Our scheme further boosts the localization efficiency by

using principal component analysis (PCA) to identify the most relevant feature vectors. The experimental results demonstrate that our system can achieve an accuracy of over 92% and an error distance smaller than 0.5m. We also investigate the effect of other parameters on the performance of our system, including packet transmission rate, the number of links as well as the number of principle components.

1.3 Contributions

To summarize, the contributions of this dissertation are as follows:

- We investigate the characteristics and advantages that FM signals can be exploited to design robust and discriminative passive system for activity recognition. Also, we conduct extensive experiments and investigate how different activities can result in the various changes of FM-radio signals. Based on these preliminary studies, we design a passive activity recognition framework, named ActiviTune, utilizing ambient FM-radio sources. To boost the recognition precision, we propose a multi-stage classification approach which can first distinguish between coarse-grained dynamic and static activities and then classify fine-grained activities. Furthermore, ActiviTune compares a variety of classification algorithms including k-Nearest-Neighbor (k-NN) and Decision Tree (DT). Also, we conduct another case study to demonstrate the proposed FM-based passive system can both localize a subject and learn his or her behavior.

- To convince our framework are capable of enabling a variety of applications to acquire user's context, we investigate the classification of FM-radio signal fluctuation for the monitoring of attention by individuals in motion towards a static object. In particular, we distinguish in a corridor, whether it is empty or populated by moving or standing individuals as well as the attention of these subjects towards poster frames in that corridor. We consider the distinction in front of which poster these subjects are walking or

standing as well as their walking speeds or changes therein. This information can provide some hint whether a person is paying attention to a specific poster in this corridor as well as the location of the particular poster.

- As stated in previous section, the user's location information is another popular context to learn. In this thesis, we design an indoor localization system which adopts the Channel State Information (CSI) measurements to mitigate the biases induced by multipath effect of RF signals in indoor spaces and hence improve the performance of localization. Furthermore, our system reduces the computation complexity to choose the most relevant feature vectors using Principal Component Analysis (PCA). To enable real-time tracking of user's trajectory, we propose Bayesian filtering schemes. We conduct experiments in four various typical indoor spaces and compare the performance of our solution with other localization schemes including *PC-DfP*, *Nuzzer* and *Pilot*. Experimental results validate the efficiency of our proposed system.

1.4 Organization of this Dissertation

The rest of this thesis is structured as follows. We discuss the background and related work in the Chapter 2. In Chapter 3, we discuss the feasibility and preliminary study of FM-radio signals based activity recognition system. Then, we present ActiviTune, a multi-stage passive activity recognition framework from ambient FM radios. Using the proposed framework, we design a context-aware application, which can implicitly monitor the attention of an user passing by a poster display. Next, we describe an indoor positioning and tracking technique to learn the user's 'location' context information in Chapter 4. Chapter 5 summarize the results and discuss the future direction.

Chapter 2

Related Work

2.1 Overview

To our understanding, it was Mark Weiser who pioneered the research area of *ubiquitous computing* in his seminal paper of 1991 [88]. Consequently, context-aware computing has become one of the publicly most visible results of Ubicomp research. In their paper "Towards a Better Understanding of Context and Context-Awareness" [2], Dey and Abowd explicitly list location and activity information as two of the four primary data categories that contribute to a user's context — alongside with time and identity. Thus, as a user's activity and location became fundamental for many Ubicomp applications, research has been focused more deeply in the fields of activity recognition [5, 6, 9] and of course location awareness [53, 91].

In the indoor environment, technologies leveraging RF signals for mining the location or activity information of an entity are recently been studied by the research field. The first section of the chapter presents the many different indoor localization systems. Then, in the section, we list many solutions for activity recognition using RF sources.

2.2 Indoor Localization Using RF

Recently, location estimation in indoor environments has gained a great deal of attention by researchers, due to the increasing demand of location-based services and applications. While a large range of information sources have been utilized, such as video [31], magnetometers [39], and magnetic resonant coupling [52], all of these systems suffer from the system cost and installation effort. In this study, of particular interest are radio frequency (RF) based solutions, as RF-channel information is a ubiquitously available source, thereby mitigating installation cost. Many systems using RF signals have been designed for precise indoor positioning for traditional *device-bound* solutions, where the target is an RF-transceiver equipped subject. Also, without requiring a person to carry any RF-emitting device, some *device-free* schemes are proposed where RF devices are placed in an area of interest to detect whether it is populated by a subject and further to identify the coordinate of him or her. We detail these two categories of localization systems following. Furthermore, we survey a plethora of methods for tracking the trajectory of a moving person using RF sources.

2.2.1 Device-bound Localization

Device-bound localization systems deal with the problem of positioning an entity equipped with an RF-emitting device. Researchers working in this direction propose to use different wireless techniques ranging from Infrared (IR) [87], Ultrasonic [29], RFID [44], Bluetooth [20] or WiFi [4]. Among these, WiFi signals are most widely adopted due to popularity and low cost. For example, as described in [4], the proposed RADAR system first builds an a priori fingerprint map by gathering the WiFi RSS measurements at different locations in the training phase, and then deduces the location by minimizing the Euclidean distance between online RSS measurements and corresponding measurements in the radio map during the test phase. The accuracy can be further improved by using CSI measurements from revised commodity WiFi devices and clustering techniques for localization [62], where experimental

evaluation asserts error distances smaller than $0.5m$. Recently, Xie et al. [96] release a new tool, *Splicer*, which can measure the CSIs from a much wider spectrum band and adjust the errors of amplitude and phase of the gathered CSIs, further improving the precision of indoor localization.

2.2.2 Device-free Localization

The assumption that a device is always carried by a subject is not realistic. A device-free system was first introduced by Youssef et al. [101] for the localization of a non-equipped entity. In recent years, various RF device-free localization schemes have been proposed [30, 91, 92, 95, 97, 101]. These schemes mainly constitute *fingerprint-based* and *model-based* solutions. We focus on the most relevant systems below.

Fingerprint-based indoor localization systems require to construct an offline radio map, and then compare it with the collected online measurements to estimate the location of the targeted person [60, 95, 97, 101]. Seifeldin et al. [60] present an RSS-based large-scale device-free localization system by analyzing the human-induced RSS changes with probabilistic techniques. Furthermore, Xu et al. [97] employ linear discriminant analysis from RSS to recognize different cells of locations, which significantly mitigates NLoS effects and thereby achieves better accuracy in cluttered indoor environments. With the evolution of WiFi PHY layer technique, Xiao et al. [95] initially estimate the position of a passive entity from CSI-related patterns. The authors propose the *Pilot* system to utilize correlations of CSI from 802.11n wireless cards as the discriminant feature to determine the location of passive subjects by a two-stage detection approach. Since CSI measurements provide more accurate channel information, the reported localization precision is dramatically increased compared to that achieved by other state-of-art RSS-based schemes. Furthermore, to ease the effort of building a radio map, a transferring positioning model based device-free system is proposed in [46], which can collect training measurements in a certain indoor environment and apply them into

other different indoor spaces with the aid of floor plans. Also, some other studies try to construct radio maps using simultaneous localization and mapping (SLAM) approaches [21, 28].

In terms of model-based localization systems, an approximate model is required upon which the localization algorithms can map the changes of RF measurements to the location of a targeted person [30, 91, 92]. For instance, Wilson and Patwari utilize Radio Tomographic Imaging (RTI) on the two-way RSS variance [91] or RSS mean fluctuations [92] between nodes arranged in a rectangle surrounding the monitored area for robust localization. To reduce the density of RF nodes, they develop a novel solution using RF transceivers in motion [30].

2.2.3 Device-free Tracking

Regarding mobility tracking systems, a generic approach is to exploit the multiple measurements in time series to reduce the positioning errors iteratively [19, 107]. In terms of device-free tracking systems, several solutions have been proposed [8, 91, 106]. For instance, in [91, 106], the authors exploit the Kalman filter for tracking a single person using the location results from variance-based RTI and subspace variance-based radio tomography respectively. Furthermore, they extend their work to track multiple persons [8]. Instead of using Kalman filter, the problem of tracking multiple persons is formulated as a data assignment problem (DAP) and solved by minimizing the total cost of DAP.

2.3 Activity Recognition Using RF

Recently, some authors also consider the detection of further situations or even activities that exceed the mere localization of subjects. Ding et al. utilised the RF-noise emitted by electric components in contemporary automobiles to classify the traffic situation in front of a traffic light [18]. The authors achieved an accuracy of more than 0.95 in most of the cases with a

single SDR employed using the mean, standard deviation, root of the mean squared (RMS) and Fast Fourier Transformation (FFT) amplitude of a signal received on a frequency range of 2.4 GHz.

The first study to report activity detection from passive entities was presented in 2011 by Scholz et al. [59]. The authors describe a system to detect walking, talking on a mobile phone and the state of the door in a typical office room with two SDR nodes (transmitter and receiver) at 900 MHz placed on both sides of the door. Walking was detected by the number of peak-to-peak amplitudes greater than a trained threshold. The door context was triggered by a static change of amplitude in the signal. In order to detect a phone call, a predefined area of the frequency spectrum was searched for a significant signal peak. The reported accuracy of the algorithm was on average above 0.8 for walking and above 0.9 for the other cases. Sigg et al. considered the detection of the activities 'sitting', 'walking' and 'standing' in a similar setting-based on the RMS, Signal-to-Noise Ratio (SNR) and Average Magnitude Squared (AMS) of the Received Signal Strength Indicator (RSSI) with an accuracy of 0.6 to 0.7 [55, 56, 72]. The authors installed in different environments two or three SDRs, from which one was used to transmit a continuous signal at 900 MHz or 2.4 GHz. With the software radios a higher sampling frequency than in previous studies is possible and also the actual channel can be sampled instead of only tracking the RSSI. Patwari et al. design a non-intrusive respiration rate monitoring system by finding the RSSI periodic change induced by the inhale of an individual. With this system, the breath rate can be estimated with a mean error of 0.03 breaths per minute.

In addition to the coarse RSSI information, later work has leveraged finer channel state information (CSI) for activity detection purpose. The channel response information on each subcarrier includes both amplitude and phase which can be utilized to provide a lot more information than RSSI readings. The E-eyes system [86] combines WiFi 2.4 GHz links from different devices (e.g., access points, thermostats, laptops) and collects fine-grained CSI mea-

surements as location-activity profiles. Utilizing a MIMO-OFDM receiver, the WiSee system [54] can distinguish 9 pre-defined gestures from different people simultaneously with accuracy of 94%. In WiVi [3], a single antenna receiver is used, while a preamble transmission stage is designed to isolate the time-varying reflections induced by the human body and null direct and wall-reflected disturbance. The system tracks the direction of the moving object using inverse synthetic aperture radar (ISAR): consecutive CSI measurements are collected over time to emulate an antenna array at the receiver.

These studies mentioned above consider a transmitter as part of and under the control of the recognition system. We recently demonstrated that a recognition is also possible when an ambient signal source is utilized. In these preliminary studies on passive device-free situation awareness we analyzed fluctuation in ambient signals from an FM radio station not under the control of the recognition system. Static environmental changes such as opened doors have been detected with an accuracy of about 0.9 [65] and a first study on suitable features to detect human activities could achieve an accuracy of about 0.8 with a two staged recognition approach [67].

DFAR is still a mostly unexplored field. Open research questions regard the optimum frequencies and the impact of the frequency on the classification accuracy, the optimum sampling rate of the signal, the detection range and the impact of this distance on the classification accuracy as well as the minimum Signal-to-Noise Ratio (SNR). Furthermore, a set of activities that can be recognized by RF-based classification is yet to be identified as well as a suitable design of the detection system. In particular, the impact of the count and height of transmitting and receiving nodes has not yet been considered comprehensively as well as even the actual necessity of a transmit node as part of the recognition system since potentially the system might utilize ambient radio. Also, it is not clear whether and how activities of multiple persons can be identified simultaneously and if features exist that enable ad-hoc DFAR systems which can be applied in novel environments without prior training. A more detailed discussion of most

of these aspects is given by Scholz et al. in [58].

Chapter 3

Mining Location and Activity Contexts from FM Sources

3.1 Introduction

In activity recognition, sensing technologies frequently require prior installation and repeated calibration within a changing environment. Examples are sensors for motion detection such as acceleration-type sensors [5, 84] as well as video or RFID-based systems [45, 89]. Clearly, the installation effort for acceleration sensors might be reduced, for instance, by proper integration of sensors in everyday objects, such as mobile phones [6, 74] or clothing [34, 37]. This has the potential to seamlessly foster the distribution of sensing equipment but also implicates the possibility of inaccurate placement, and therefore increases the measurement noise or the requirement for frequent sensor calibration [24]. Video, on the other hand, is mostly restricted to indoor environments since it is well limited in range and is heavily impacted in its accuracy by changing light conditions [81]. However, there are other ubiquitously available, accurate and rich information sources from which we can detect activities of individuals indoors and outdoors and which come preinstalled with their fixed infrastructure [9, 25, 49, 82]. Some of these information sources do not require sensors to be attached to the sensed entities which

further implicitly reduces the effort required for calibration.

The RF-channel for wireless communication might constitute such information source. Nearly all contemporary electronic devices contain some kind of interface to the RF-channel. With the upcoming Internet of Things (IoT) in which virtually all objects are enhanced by communication means, this situation will still greatly intensify [26, 33]. Furthermore, since the available wireless spectrum is sparse [48], we are always surrounded by some sort of electromagnetic waves transporting signals associated, for instance, to audio (FM, AM), video (DVB-T), Speech (GSM, UMTS) or data (LTE, Wimax, HSDPA, Bluetooth, Wifi, ZigBee). The RF-channel was recently utilized for the localization of passive individuals [51]. The localization accuracy reached 0.5 meters for a single subject [35] or about 1 meter for up to five subjects simultaneously [103]. Since such high accuracy was achieved, it might be feasible that also other classes classes than activity and presence are detectable. The classification of simple activities is the natural extension of the detection of location to further classes.

We propose to use the RF-channel as an information source to detect activities of individuals by monitoring peculiarities of RF-based features induced by these very activities. The contributions of this chapter are

a) a study of features from the RF-channel which are suitable for the recognition of activities;

b) a 2-staged passive device-free activity recognition (DFAR) system (ActiviTune) which distinguishes between static and dynamic activities in order to classify ‘empty room’, ‘standing’, ‘lying’, ‘crawling’ and ‘walking’;

c) a case study with three subjects in which we classify the above activities in two scenarios;

d) studies on the impact of the distance and count of receiver devices on the classification accuracy of FM-based passive DFAR systems;

e) the direct comparison of classification accuracies achieved by FM radio-based passive device-free activity recognition, software defined radio (SDR) and Wireless Sensor Node (WSN)-based active device-free activity recognition and accelerometer-based activity recognition.

f) a further experiment to demonstrate that our system can be extended not only to detect activities but also to recognize the area where a particular activity occurs.

g) a use case where the proposed FM-based passive DFAR system can be leveraged to succeed in monitoring the attention of a subject in front of a poster frame.

The rest of this chapter is structured as follows. Section 3.2 discusses the theoretical foundation and feasibility of our recognition system, and describes the implementation of the passive transmitter-free system-based on the RF-channel. The implementation of the recognition system is discussed in section 3.3 and results on the recognition performance are presented in section 3.4. We detail the classification accuracy for one-stage and two-stage approaches (section 3.4.1), consider the impact of the distance to the receiver device (section 3.4.2), show similar classification accuracies in an alternative environment (section 3.4.3), compare the accuracy reached to a passive WSN-based DFAR system (section 3.4.5) and to accelerometer-based activity recognition (section 3.4.4) as well as demonstrate the possibility to jointly detect both the location and activity of an individual with two synchronized USRP devices (section 3.4.6). Section 3.5 presents a use case of attention monitoring with the proposed FM-based passive DFAR system. Finally, section 3.6 draws our conclusion and discusses some open issues for future work.

3.2 Background and Feasibility Study

Currently, virtually all electronic devices provide an interface to the radio spectrum. However, apart from the data modulated onto the carrier, other available information, such as the signal

strength, signal to noise ratio (SNR) and energy in the signal frequency domain is usually discarded. In the following sections, we discuss the use of this additional information broadcast by an ambient transmitter not under the control of the system for activity recognition. Furthermore, we discuss various signal sources in different spectrum bands, such as FM radio stations, WiFi access points and GSM/UMTS/LTE base stations which can all be leveraged for passive RF-based DFAR systems. From these, we identify several advantages of FM signals compared with other RF signals.

3.2.1 Principles of RF signal propagation

RF signals are electromagnetic waves whose physical attributes are defined by their amplitude, phase offset and frequency. Assume a signal observed at a receiver, at some frequency f_c [Hz]. Naturally, signal propagation is roughly omnidirectional. In the event that radio waves encounter any concrete structure such as an object or individual, the corresponding signal components will be damped (continue their path with reduced energy) or even completely blocked. Additionally, the signal can be reflected or scattered at this event. The signal $\zeta_{\text{rec}}(t)$ at a receiver is composed of distinct signal components that arrive along various signal propagation paths between a transmitter and a receiver. All these signal components for a message $m(t)$ add up at the receiver to form a superimposition

$$\zeta_{\text{rec}}(t) = \Re\left[\left(m(t)e^{j2\pi f_c t}\sum_{i=1}^{n}\text{RSS}_i e^{j\gamma_i}\right) + \zeta_{\text{noise}}(t)\right] \tag{3.1}$$

In equation (3.1) the RSS_i denotes the received signal strength of the i-th signal component out of n received signal components. The value γ_i accounts for the phase offset in the received signal components $i \in \{1, \ldots, n\}$ due distinct signal propagation times. Due to signal reflection at objects, a change in the environment such as the presence of persons or changed location of objects will result in differing signal propagation paths between a transmitter and a receiver.

Such a change might then alter the signal strength at the receiver and also the phase of the signal $\zeta_{\text{rec}}(t)$. Furthermore, the noise figure $\zeta_{\text{noise}}(t)$ received together with $\zeta_{\text{rec}}(t)$ might change depending on the environment. Straightforward examples are indoor and outdoor environments which will likely differ in their noise and signal-to-noise level.

3.2.2 Discussion of various RF-technologies for passive DFAR

In the literature, RF-sensing for localization or situation detection is applied on various signal frequencies and technologies such as WiFi, GSM or FM-radio [36, 42, 61, 102, 105]. For activity recognition, we believe that FM radio is currently best suited for RF-based DFAR systems for the following reasons.

As shown in [13, 53], leveraging broadcasted FM radio signals can augment the accuracy of indoor localization, due to signal characteristics of FM radio, such as the low operating frequency range, the simple modulation mechanism and the wide area of coverage. It was demonstrated that these characteristics can be exploited to design more robust and discriminative signatures for RF-fingerprinting than WiFi [13, 53] and GSM [53].

FM radio signals experience, when compared with WiFi, 3G or 4G signals, lower variation in signal strength over time [13]. Since we employ signal strength variation also for our features to distinguish activities, FM radio signals induce a lower process noise than signals from WiFi, 3G or 4G systems. For passive DFAR systems utilizing ambient signal sources, also the sensitivity to changing weather conditions must be considered in the choice of alternatives. FM radio is, compared to the other named systems, which operate at higher frequencies, less susceptible to weather conditions, such as rain and fog [83]. Additionally, in order to increase spectrum efficiency, spread spectrum techniques such as frequency hopping or code divisioning are employed in WiFi, 3G and 4G access points. In particular, participating devices that follow these schemes are required to closely comply to the parameters provided by the base station. Hopping schemes or CDMA codes are typically not disclosed to exter-

nal devices. Since we assume that the ambient signal source is not under the control of the activity recognition system, it is natural to assume that the receiver is not capable of following any of these spread spectrum schemes. This likely results in significant variation in the observed signal strength or also in very low signal strengths in a given frequency band. For activity recognition, these spread spectrum techniques therefore increase process noise which makes a successful utilization of the corresponding signals for passive DFAR systems more challenging. The modulation applied to FM radio is much simpler so that it is easier for a third party receiver to monitor the evolution of the signal. Furthermore, FM radio stations are widely implemented and continuously broadcast signals with higher coverage than WiFi, 3G or 4G systems. A passive DFAR system utilizing FM radio can therefore be deployed virtually anywhere without considering the presence of surrounding implicit RF radio sources. Finally, FM radio is embedded in many contemporary electronic devices. For the above reasons, we believe that FM radio is best suited for the utilization in a passive DFAR system.

3.2.3 Environment induced characteristics of RF signals

As mentioned above, objects encountered by the electromagnetic waves during signal propagation can result in the alteration of physical properties of the received signal. Assume an ambient transmitter emitting signals. When the environment changes, also the paths of the distinct signal components that make up ζ_{rec} change. However, the impact of the received signal is bigger when the environmental changes occur in the proximity of the receiver. Environmental changes farther away from the receiver constitute in general noise and have only minor effect on the received signal as the modified signals are greatly damped before they reach the receiver. Modifications in the proximity of the receiver, however, have greater effect on the received signal. For instance, assume that several individuals block some of the signal paths in an environment with their body. While individuals are moving, the signal paths are constantly altered, resulting in a higher fluctuation of the received RF signal. We assert that

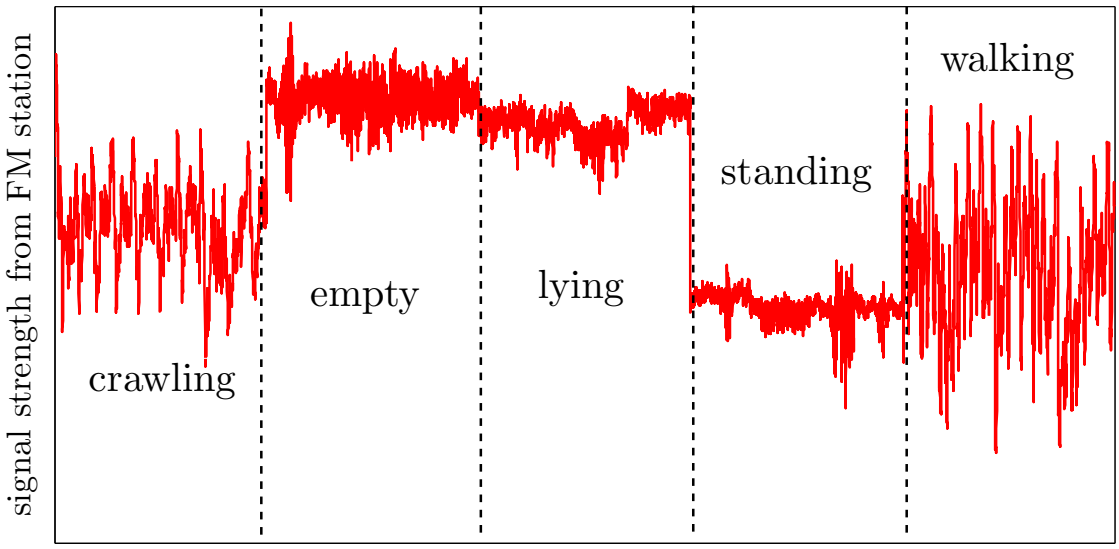


Figure 3.1: Evolution of signal strength for all activities considered, performed by a single subject

the presence and movement of individuals in a room induces characteristic patterns on the time-evolution of a received signal. Furthermore, we believe that different activities induce different and detectable patterns. This property was utilized also for the device-free localization of individuals [36, 42, 61, 102, 105]. In contrast to these studies, we refrain from employing a dedicated transmitter as part of the recognition system and distinguish activities instead of locations. Figure 3.1 illustrates the received signal strengths from an ambient FM radio station over 1 minute for various activities of a single subject in an indoor environment.

The figure shows a correlation between the characteristics of an RF signal and the activities conducted. This observation suggests that with proper features and classification schemes it is possible to distinguish the activities considered from ambient signals. Over time, we experienced only small changes in the signal strength of the received FM signal. In particular, over a sampling period of about one hour, the offset between the maximum and minimum average value is lower than $2 \cdot 10^{-5}$, while the average signal strength observed was about $3.5 \cdot 10^{-3}$.

3.2.4 An application scenario for DFAR

Most accidents happen at home. The primary reason for these accidents are falls which make up about 40 percent of the total number of accidents. Most of these accidents leave the affected person in an unusual posture such as lying at an unusual location. While the automatic detection of fall and fall prevention has gained large interest in the research community and various approaches have been proposed, these alarm system either need body-attached sensors, require the installation of a complex infrastructure or have strong privacy related implications as, for instance, video based systems. By utilising the RF-sensor for this kind of detection we would reduce privacy issues, avoid the need of having to carry sensors and ideally reduce installation requirements to a minimum. The sensor could further become a crucial component of (Health) Smart Home systems relieving users from the necessity to wear a device. In fact, for Smart Home systems, the sensor needs to provide a rough localisation capability as well as the recognition of at least a basic set of activities of daily living. Among such activities are walking, standing and sleeping. Considering the demographic change in developing and developed countries, the application of the RF-sensor for alarm systems or Smart Homes could further play an important role towards the extension of self-sustained living of the elderly. The present study illustrates the potential of the ubiquitously available RF-sensor for the detection of relevant activities in Smart Home environments.

3.3 FM-based Passive Activity Recognition Framework

In the following sections we propose a passive DFAR approach based on environmental signals sampled from an FM radio station. In particular, we introduce the recognition system in section 3.3.1 and discuss possible features in section 3.3.2.

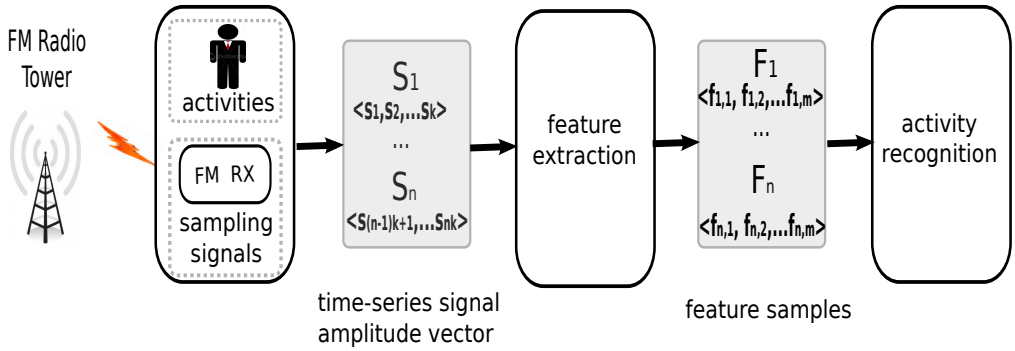


Figure 3.2: An overview of the signal acquisition, processing and activity recognition of the passive DFAR system

3.3.1 A passive DFAR system

We propose the passive FM-based DFAR system sketched in Figure 3.2.

In our approach we utilize signals continuously broadcast by an FM radio station. A subject in proximity of a receive antenna can by her movement and activities change the set and properties of signal components that constitute the received signal $\zeta_{\text{rec}}(t)$. In the figure, the receiver is labelled FM RX. The receiver analyses the evolution of the measured signal over time and groups series of continuous measurements $s_1, s_2, \ldots, s_t$ in sample windows $S_1, \ldots, S_n$. Each S_i consists of k consecutive samples $\langle s_{(i-1)k+1}, s_{(i-1)k+2}, \ldots, s_{ik} \rangle$. From these vectors S_i, sets of features $\mathrm{F}_i = \langle f_{i,1}, f_{i,2}, \ldots, f_{i,m} \rangle$ are extracted. These features are then used as a basis for the recognition of activities.

In our case studies, we randomly divided the collection F of all features F_i into a training set Tr and a classification set Cl that meet the conditions $\mathrm{Tr} \cup \mathrm{Cl} = \mathrm{F}$ and $\mathrm{Tr} \cap \mathrm{Cl} = \emptyset$. The set Tr is used to train the classifiers. After the training, classifiers will process Cl.

For a set of k activities $\mathcal{A} = \{a_1, \ldots, a_k\}$ let $\mathcal{I}(a_i)$, with $\mathcal{I}(a_1) \cup \mathcal{I}(a_2) \cdots \cup \mathcal{I}(a_k) = \mathrm{Cl}$, be the total number of instances for activity a_i and $\mathcal{I}_{\text{cor}}(a_i)$ be the number of correctly classified instances for this activity in which the classification matches the ground truth. We define the

accuracy by which an activity a_i can be detected as

$$\mathcal{ACC}(a_i) = \frac{\mathcal{I}_{\text{cor}}(a_i)}{\mathcal{I}(a_i)} \tag{3.2}$$

3.3.2 Features for passive DFAR

Extracting and interpreting suitable features is essential for the recognition accuracy. For RSSI-based RF-samples the mean and variance have been successfully used [102] while for activity recognition in general good results have been reported for the mean, energy and entropy [5]. Observe that contemporary wireless devices commonly provide an interface to the signal amplitude for higher layer applications. In particular, the RSSI is commonly provided by contemporary transceiver hardware. Therefore, the features presented here are similarly applicable for applications on such mobile devices. In our approach, the features we utilized are the mean Avg_i and variance Var_i of S_i, the energy E_i in the frequency-domain and the entropy T_i of S_i. That is, the corresponding characteristic instance, F_i, is represented by the vector$\langle \text{Avg}_i, \text{Var}_i, \text{E}_i, \text{T}_i \rangle$.

mean value over a single window reflects the average signal strength over that period and thus tracks the evolution of signal strength over time. The mean for the vector S_i can be computed as

$$\text{Avg}_i = \left(\frac{\sum_{t=(i-1)n+1}^{in} \text{s}(t)}{n} \right)^2. \tag{3.3}$$

The variance of a window of signal strength measurements is frequently utilised as a reference value which can represent how characteristically the signal's strength changes within a certain time interval. For the interval S_i, the variance is calculated as

$$\text{Var}_i = \sqrt{\frac{\sum_{t=(i-1)n+1}^{in} (s(t) - \text{Avg}_i)^2}{n}}. \tag{3.4}$$

Frequency domain features have been widely used to capture the periodic nature of activities,

such as walking, running and cycling [5]. We use two frequency-domain features, the energy and entropy of S_i which have shown good results in previous work [5, 23]. To compute the normalized spectral energy, the first step is to transform the amplitude from the time domain into the spectral values in frequency domain using the FFT

$$\mathrm{FFT}_i(k) = \sum_{t=(i-1)n+1}^{in} \mathrm{s}(t) e^{-j\frac{2\pi}{N}kt}. \tag{3.5}$$

In equation (3.5) the value $\mathrm{FFT}_i(k)$ denotes the kth frequency component of S_i. The normalized energy of a signal can be derived as the squared sum of its probability density of spectrum in each frame. The probability of each spectral $\mathrm{FFT}_i(k)$ band is computed as

$$\mathrm{P}_i(k) = \frac{\mathrm{FFT}_i(k)^2}{\sum_{j=1}^{n} \mathrm{FFT}_i(j)^2}. \tag{3.6}$$

Hence, we calculate the normalized spectral energy as

$$\mathrm{E}_i = \sum_{k=1}^{n} \mathrm{P}_i(k)^2. \tag{3.7}$$

The metric of the information entropy in the frequency domain can be used to discriminate activities which have similar energy values. The following equation defines the normalized spectral entropy

$$\mathrm{T}_i = -\sum_{k=1}^{n} \mathrm{P}_i(k) \log_2 \mathrm{P}(k). \tag{3.8}$$

3.4 Experimental Evaluation

To demonstrate the viability of the passive transmitter-free system, we performed experiments in an indoor environment depicted in figure 3.3 with 3 participants to distinguish the five situations or activities empty room, lying, standing, walking and crawling. The semi-

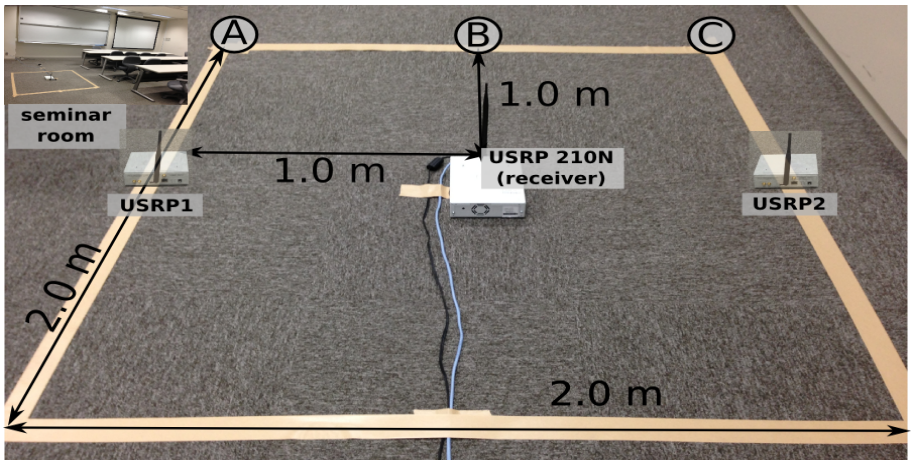


Figure 3.3: Schematic illustration of the seminar room used in our case study

nar room ($8.5m \times 6.5m$) is equipped with tables, chairs, a white board and some electronic equipment (e.g. a television and a projector). The activities are recognised within a $2m \times 2m$ recognition area in section 3.4.1 and the impact of the distance in which the activity is performed is considered in section 3.4.2. Section 3.4.3 discusses the application in an alternative environment and section 3.4.4 investigates the accuracy achieved with accelerometers for comparison.

3.4.1 Recognition of activities from FM signals

We use a $2m \times 2m$ square region as the recognition area in the seminar room with a receiver in its center. The receiver utilizes the WBX[1] daughter board incorporated with a USRP N210[2] SDR device. The antenna utilized is the VERT900[3] model with 3dBi antenna gain. In figure 3.3 we have marked the locations at which activities were performed with A, B and C (cf. table 3.1). To collect sufficient data footage, every activity was conducted approximately 3 minutes while the transceiver tuned to a FM radio station continuously recorded the signal strength with a sample rate of R = $64Hz$.

[1] https://www.ettus.com/product/details/WBX
[2] https://www.ettus.com/content/files/2987_Ettus_N200-210_DS_FINAL_1.27.12.pdf
[3] https://www.ettus.com/product/details/VERT900

Activity	Location	Description
Stand	B	Standing still on location A
Lying	A-C	Lying across location A to C
Walk	All	Walking around freely in the 2×2 area
Crawl	All	Crawling around freely in the 2×2 area
Empty	–	Empty room without movement or individuals.

Table 3.1: Activities performed with respect to locations shown in figure 3.3.

We used 128 samples for each window S_i, spanning a total of 2 seconds. After the feature extraction, the Orange data mining Toolkit[4] was utilized to classify the set of instances. A decision tree (DT) and a k-nearest-neighbour classifier (k-NN) were leveraged for testing the classification performance. The k-NN classifier utilized 5 neighbours and weights their distance by the Euclidean distance. The decision tree utilizes at minimum 10 instances in its leaves for pre-pruning and a recursive merge of leaves of the same major class with an m-estimate of 2 for post pruning.

The accuracy of the classification is presented in table 3.2. Figures in brackets denote the standard deviation over the 10 test cases. Table fields with very low values (i.e. 0.0(0.0)) are left blank to improve readability.

We observe that all activities can be correctly recognized with a mean accuracy of 0.69 and 0.747 by the Decision Tree and k-NN classifiers respectively. Also, based on the experimental results, we can conclude that the dynamic activities generally are less likely to be correctly classified than the static ones.

A two-stage classification scheme has demonstrated to be superior in previous work [7, 80] compared with a one-stage method. In [80], it was first distinguished whether a dynamic or static activity was observed in the first stage. As for the next stage, the instance recognized as dynamic would be further distinguished to be either walking or driving. In our case, the

[4] http://orange.biolab.si/

Table 3.2: Mean accuracy for one-stage method of the activities 'lying', 'standing', 'crawling', 'walking' and 'empty' from amplitude based features extracted from an FM-radio signal in the environment depicted in figure 3.3

(a) Classification accuracy achieved by a k-NN classifier

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**.739**(.324)	.087(.028)	.145(.044)	.029(.005)	(.001)
	lying		**.733**(.276)	.213(.155)	.027(.004)	.027(.004)
	stand		.157(.039)	**.843**(.187)	.0(.001)	.0(.001)
	walk		.083(.021)	.012(.002)	**.706**(.336)	.2(.074)
	crawl		.054(.016)		.230(.107)	**.716**(.288)

(b) Classification accuracy achieved by a DT classifier

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**.780**(.142)	.0(.001)	.220(.019)	.0(.002)	
	lying	.0(.001)	**.507**(.419)	.440(.338)	.053(.006)	
	stand		.214(.11)	**.786**(.208)		
	walk	.035(.005)	.047(.006)	.012(.003)	**.659**(.361)	.247(.157)
	crawl	.013(.003)	.041(.009)		.230(.108)	**.716**(.293)

five activities our system recognizes can be grouped into two categories: stationary (including empty room, lying and standing) and dynamic (including walking and crawling). Figure 3.4 indicates that it might be easy to distinguish between stationary and dynamic activities in order to reduce the number of distinct classes for the actual classification. In comparison to the stationary activities, the volatility of fluctuation of the signal strength of dynamic activities is much more significant. Instances are recognized by DT and k-NN classifiers. Results after 10-fold cross validation are shown in table 3.3 are also average accuracies, standard deviations are written in brackets.

Table 3.4 depicts the overall accuracies and standard deviation for one and two stage classifiers respectively. For both classifiers the accuracy can be remarkably improved with the two-stage approach with an increase to 0.84 and 0.856 and a 15% and 11% gain for Decision Tree and k-NN classifiers. Furthermore, observing the confusion matrices depicted in

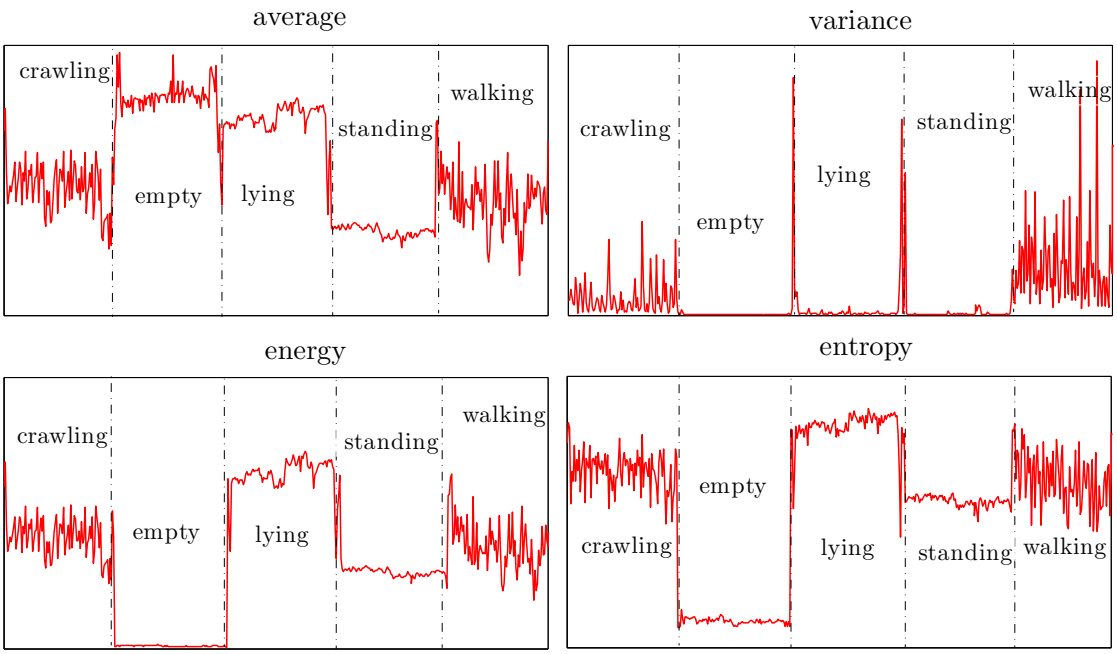


Figure 3.4: Features over 128 samples for 5 activities utilizing the raw signal strengths shown in fig 3.1

tables 3.2 and 3.3, the performance of the two-stage method is superior for each of the five activities. The reason for this is the high accuracy for the first stage and the reduced activity space for the second stage.

3.4.2 Influence of distance

The distance to the receiver in which an activity is performed impacts the intensity of the characteristic pattern induced on the RF-channel and therefore also the classification accuracy. We investigate the impact of the distance between the location where the activities were performed and the receiver. In the lecture room shown in figure 3.3 all activities have been previously performed within a recognition area of 2 m^2.

In order to estimate the impact of distance, we again conduct activities in the same setting. The same activities are conducted by three subjects in increasing distance to the receiver. In particular, each subject repeated the activities in 1, 2, 3, 4, 5 and 6 meters distance. The accuracies achieved over 10 two-stage classifications with k-NN and DT classifiers are shown

Table 3.3: Mean accuracy for two-stage method of the activities 'lying', 'standing', 'crawling', 'walking' and 'empty' from amplitude based features extracted from an FM-radio signal in the environment depicted in figure 3.3

(a) Classification accuracy achieved by a k-NN classifier

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**.89(.053)**	.089(.021)	.007(.002)		.011(.002)
	lying	.077(.019)	**.87(.057)**	.053(.017)		.004(.001)
	stand	.053(.013)		**.94(.034)**		.008(.002)
	walk	.073(.009)	.051(.011)	.051(.008)	**.74(.076)**	.080(.004)
	crawl	.051(.011)	.005(.001)	.1(.039)		**.84(.063)**

(b) Classification accuracy achieved by a DT classifier

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**.83(.102)**	.085(.036)	.070(.019)		.011(.005)
	lying	.042(.007)	**.89(.031)**	.056(.009)		.004(.001)
	stand	.034(.007)	.004(.001)	**.95(.010)**		.008(.003)
	walk	.089(.026)	.073(.024)	.012(.003)	**.75(.133)**	.173(.082)
	crawl	.037(.008)	.009(.002)	.107(.034)	.069(.022)	**.78(.110)**

in figure 3.5.

Table 3.5 shows the confusion matrices for each distance.

3.4.3 Alternative environments

We have demonstrated an FM-based passive DFAR system in a single environment. Clearly, the environment conditions and reception characteristics might impact the classification accuracy and even the feasibility of FM-based passive DFAR. To validate our system and demon-

Table 3.4: Activities performed with respect to locations shown in figure 3.3.

	accuracy	**standard Deviation**
one stage by DT	0.690	0.215
two stage by DT	0.840	0.162
one stage by k-NN	0.747	0.231
two stage by k-NN	0.856	0.172

Table 3.5: Mean accuracy for two-stage method of the activities ‘lying’, ‘standing’, ‘crawling’, ‘walking’ and ‘empty’ from amplitude based features extracted from an FM-radio signal in the environment depicted in figure 3.3

(a) Confusion matrix for the k-NN classifier at location L_1

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**1.0(.0)**				
	lying	.083(.021)	**.917(.072)**			
	stand			**.938(.074)**	.033(.006)	.029(.005)
	walk	.028(.003)		.033(.003)	**.889(.065)**	.050(.006)
	crawl			.062(.006)	.081(.010)	**.857(.089)**

(b) Confusion matrix for the k-NN classifier at location L_2

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**.889(.085)**				.111(.008)
	lying		**1.0**			
	stand	.056(.007)		**.916(.063)**	.028(.005)	
	walk	.029(.005)		.086(.013)	**.829(.166)**	.057(.008)
	crawl	.108(.021)	.054(.008)	.054(.008)	.054(.008)	**.730(.183)**

(c) Confusion matrix for the k-NN classifier at location L_3

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**1.0**				
	lying		**.914(.052)**			.086(.019)
	stand			**.763(.204)**	.132(.032)	.105(.023)
	walk		.027(.006)	.162(.046)	**.648(.225)**	.162(.031)
	crawl		.058(.009)	.147(.032)	.088(.013)	**.706(.144)**

(d) Confusion matrix for the k-NN classifier at location L_4

		Classification				
		empty	lying	stand	walk	crawl
Truth	empty	**.972(.062)**		.027(.003)		
	lying		**.865(.081)**	.135(.023)		
	stand		.143(.012)	**.771(.087)**	.086(.009)	
	walk		.088(.013)	.324(.138)	**.588(.219)**	
	crawl	.052(.006)	.068(.008)	.103(.014)	.194(.032)	**.583(.163)**

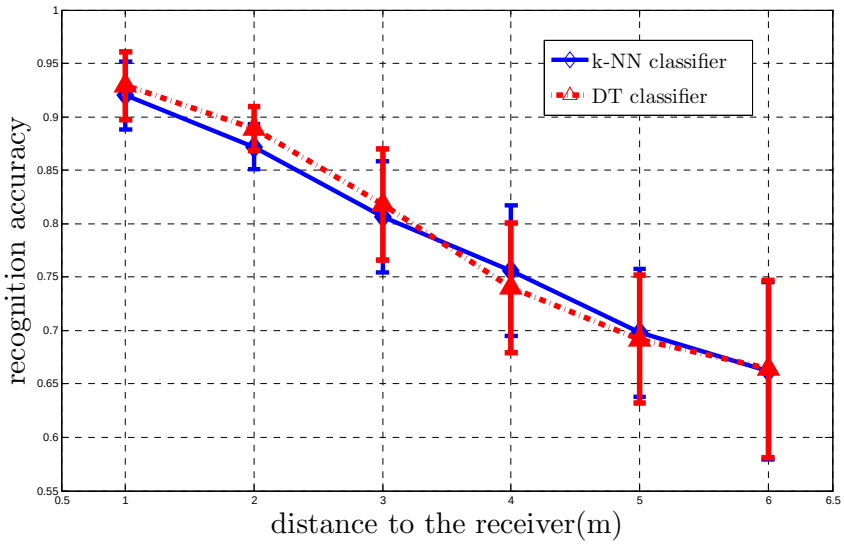


Figure 3.5: Activity recognition accuracy conducted by 3 subjects at area L_1 to L_6 by k-Nearest-Neighbor classifier

strate the adaptation to various environmental deployments, we repeated the experiments described in the previous section (five activities, three (different) subjects, 10 fold cross validation) also in another environment. The smaller seminar room we utilized has a size of approximately 5 m $\times$ 3 m and is equipped with a desk and chairs (cf. figure 3.7).

In contrast to the lecture room depicted in figure 3.3, the seminar room is located in a different floor of the building and approximately in the center of the building with no windows or connection to the outer walls. In this case, we choose the detection area as a circle with a radius of 1.25 meters. The FM-USRP receiver is located in its center. The activities of crawling and walking were conducted along the path of the circle and the other two activities, standing and lying, were performed at the location labelled S in the figure. Using the identical approach for feature extraction and two-staged activity classification we achieved overall accuracies for the five activities after 10 runs of about 0.892 and 0.914 by k-NN and DT classifiers respectively. In contrast, in the 2 m $\times$ 2 m detection area of the previous scenario depicted in Figure 3.3, an overall accuracy of 0.856 and 0.84 was reached by k-NN and DT classifiers respectively.

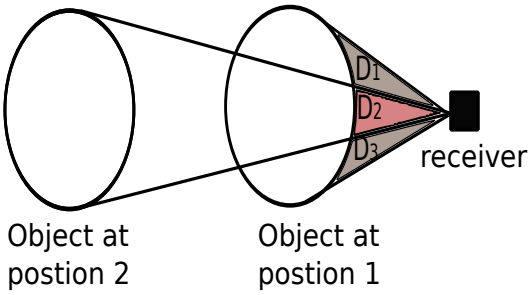


Figure 3.6: Visualization of object impact on the signals with different distance to the transceiver

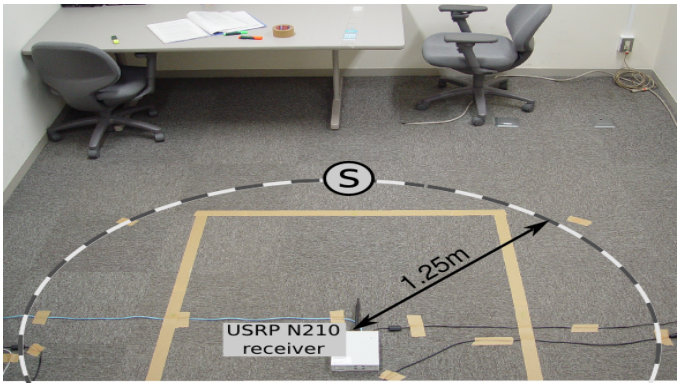


Figure 3.7: Schematic illustration of the seminar room used in the case study

We attribute the better performance of the recognition system in the seminar room to the central location of the receiver so that the distance to the locations at which the activities were performed was shorter. This experiment confirms that the FM-based passive DFAR system detailed can be employed at various locations and reaches similar accuracy with the features identified.

3.4.4 Classification accuracy with accelerometer devices

Motion and activity detection leveraging acceleration-type sensors have widely been studied [5, 40, 84] and are applied also in mature commercial applications, for instance, the

nike+iPod Sport Kit which can measure the pace of a walk or run[5], the nintendo Wii[6], Sony PlayStation 3 DualShock 3 or also for image stabilization in camcorders. Therefore, we would like to compare the recognition accuracy of FM-based passive DFAR systems with an accelerometer-based activity recognition system. Since the accelerometer is in direct body contact, we expect a better recognition accuracy. Clearly, this is achieved at the cost of having to equip subjects. We utilized an off-the-shelf accelerometer shipped with an iPhone 4 smartphone. We did not tightly calibrate the accelerometer or fix it at any specific body part but placed the phone in upright position in the right front pocket of the trousers the subjects were wearing. The accelerometer samples were taken simultaneously to the FM-based passive DFAR experiments described in section 3.4.1.

Hence, while each individual was performing the four activities, namely, standing, lying, crawling and walking, the person was equipped simultaneously with the smartphone which continuously aggregated the accelerometer data at 40 samples per second. Since the accelerometer is not capable of detecting the empty room, this class was removed when calculating the accuracy for the FM-based passive DFAR system. Similar to the experiments described above, we utilized the same features over a window size of 2 seconds. Table 3.6 depicts the accuracies for the passive FM-based and the accelerometer-based system utilizing k-NN and DT classifiers.

3.4.5 Classification accuracy of active WSN based DFAR

Similarly to the previous section, we compared the accuracy that can be achieved by a passive FM-based DFAR system to an active DFAR system in which the channel is monitored by wireless sensor network nodes. In particular, we utilised the INGA and Jbee sensor nodes.[7] [8] These nodes have no direct access to the wireless channel as the USRP devices have it. In

[5] http://www.apple.com/ipod/nike/
[6] http://wii.nintendo.de
[7] INGA: http://www.ibr.cs.tu-bs.de/projects/inga/
[8] Jbee: https://github.com/teco-kit/Jennisense/wiki/Hardware.

Table 3.6: activity recognition accuracies in comparison between RF sensor and accelerometer sensor performed in a conference room illustrated as fig 3.7

(a) Classification accuracy for FM-based passive DFAR by a k-NN algorithm

		Classification			
		lying	stand	walk	crawl
Truth	lying	**1.0**			
	stand	.056	**.98**	.022	
	walk	.023		**.874**	.102
	crawl	.044		.144	**.811**

(b) Classification accuracy for FM-based passive DFAR by a DT algorithm

		Classification			
		lying	stand	walk	crawl
Truth	lying	**1.0**			
	stand	.056	**.98**	.022	
	walk		.023	**.891**	.086
	crawl		.044	.127	**.829**

(c) Classification accuracy for accelerometer-based DFAR by a k-NN algorithm

		Classification			
		lying	stand	walk	crawl
Truth	lying	**.976**	.024		
	stand		**1.0**		
	walk			**.955**	.045
	crawl			.253	**.748**

(d) Classification accuracy for accelerometer-based DFAR by a DT algorithm

		Classification			
		lying	stand	walk	crawl
Truth	lying	**.988**	.012		
	stand	**1.0**			
	walk	.023		**.874**	.102
	crawl			.268	**.732**

particular, the measure that can be obtained by these nodes is the RSSI which is transmitted with each received packet so that the sample rate is dependent on the count of packets received. Due to the limited number of available amplification stages the resolution of RSSI (typically less than 60 distinct stages) is much lower than that of the USRP (typically a 32-bit complex float). Although we can therefore expect a reduced classification accuracy with these devices, the scenario is more realistic for an active system since SDR devices are not typically employed in common equipment. Classification accuracies for the INGA nodes, transmitting at 100 packets per second in the scenario depicted in figure 3.7 are presented in table 3.9a and table 3.9b for the four activities lying, standing, walking and crawling for k-NN and DT classifiers, respectively. We observe that the classification accuracy of the WSN-based active DFAR system is significantly lower than the accuracy achieved by the passive FM-based DFAR system (cf. table 3.9).

In a follow-up experiment, using Jbee WSN nodes, we investigated if activity recognition

Table 3.7: Class accuracies (true positive rates) achieved by a k-NN classification algorithm for an active WSN-based DFAR system when increasing the number of nodes in the experimentation area

		Number of employed WSN Nodes						
		2	3	4	5	6	7	8
Acc. k-NN	lying	.742	.920	.969	.985	.993	.998	1.0
	standing	.483	.756	.866	.916	.942	.956	.961
	walking	.530	.617	.683	.719	.735	.745	.750
	crawling	.472	.728	.852	.904	.940	.963	.993

Table 3.8: Class accuracies (true positive rates) achieved by a DT classification algorithm for an active WSN-based DFAR system when increasing the number of nodes in the experimentation area

		Number of employed WSN Nodes						
		2	3	4	5	6	7	8
Acc. DT	lying	.772	.897	.929	.950	.960	.967	.961
	standing	.500	.710	.787	.823	.837	.834	.849
	walking	.580	.684	.732	.774	.802	.812	.829
	crawling	.483	.666	.731	.768	.770	.779	.776

accuracy can be increased when employing multiple transceivers. As WSN nodes or other RSSI-capable devices are cheap or already available in a users' home, the deployment of multiple transceivers seems practical. For this experiment we installed a total of 8 transceivers in an office room of 5m $\times$ 4m while two subjects conducted the previously described activities for 2.5 minutes. Transceivers communicated using a time division multiple access (TDMA) scheme sending and receiving 40 packets per second. Computed features are mean and variance over one second. Table 3.7 and table 3.8 show the increase in the recognition accuracy with increasing number of nodes for k-NN and DT classification algorithms. Average accuracy (true positives rates) for a single link (two nodes) is lower compared to the INGA experiment as the area in which the activities are conducted is more than twice the size and since the sampling rate is chosen considerably lower to achieve a more realistic setting.

We conclude that in current environments and also in the upcoming Internet of Things, an FM-based passive DFAR system is likely to be more accurate than feasible active DFAR

Table 3.9: activity recognition accuracies in comparison between RF sensor and accelerometer sensor performed in a conference room illustrated as fig 3.7

(a) Active WSN-based DFAR (k-NN)

		Classification			
		lying	stand	walk	crawl
Truth	lying	**.882**	.118		
	stand	.12	**.869**	.007	.004
	walk			**.953**	.047
	crawl		.01	.439	**.551**

(b) Active WSN-based DFAR (DT)

		Classification			
		lying	stand	walk	crawl
Truth	lying	**.843**	.153		.004
	stand	.152	**.845**		.004
	walk		.003	**.893**	.104
	crawl	.01		.449	**.541**

installations while reaching an accuracy comparable to accelerometer based systems for the activities lying, standing, walking and crawling. The active WSN-based system could achieve higher accuracy only when a high count of nodes were employed.

3.4.6 Joint localization and activity recognition

In the event that a radio wave encounters any concrete structure such as an object or individual, the main signal component will be damped (continue its path with reduced energy) or even completely blocked. These effects also impact the received signals of USRP devices. The impact of these effects is proportional to the distance between the blocking object and the receiver. On the other hand, when a dynamic activity is performed by an individual, the amplitude of the RF signal strength will greatly fluctuate in response. We observe that whatever activities performed, the received signals differ due to the position deviation of receivers. Also, the characteristic pattern of signals will be altered by presence and movement of an individual. This observation suggests that with proper features and classification schemes it is possible to simultaneously distinguish the location and the activities of an individual considered from ambient signals.

Our system is extended not only to detect activities but also to recognize the area where a particular activity occurs. Our approach focuses on utilizing the FM signals from two synchronized USRP devices to detect activities in a particular position. To conduct the experiment, we

Table 3.10: Accuracy of joint localization and activity recognition for the activities 'lying', 'standing', 'walking', 'walking' and positions S_1 and S_2 in a corridor

(a) Classification accuracy achieved by a k-NN classifier

		Classification (%)					
		st at S_1	st at S_2	ly at S_1	ly at S_2	wa at S_1	wa at S_2
Ground truth	st at S_1	**96.92**	0	0	0	3.08	0
	st at S_2	0	**96.83**	0	0		3.17
	ly at S_1	0	8.70	**76.81**	14.49	5.63	0
	ly at S_2	0	2.82	11.27	**80.28**	5.63	0
	wa at S_1	22.22	0	0	0	**71.43**	6.35
	wa at S_2	15.39	0	0	0	18.46	**66.15**

(b) Classification accuracy achieved by a DT classifier

		Classification (%)					
		st at S_1	st at S_2	ly at S_1	ly at S_2	wa at S_1	wa at S_2
Ground truth	st at S_1	**81.54**	1.54	0	1.54	0	15.39
	st at S_2	0	**79.37**	14.29	1.59	0	4.76
	ly at S_1	0	0	**79.71**	18.84	1.45	0
	ly at S_2	5.63	0	1.41	**78.87**	14.08	0
	wa at S_1	19.05	0	12.30	0	**51.19**	17.46
	wa at S_2	7.69	7.69	0	12.31	21.54	**50.77**

place two synchronized USRP receivers along the wall of a corridor with a distance of 4 meters. In this scenario, we distinguish three activities, namely, 'standing', 'lying' and 'walking' in two 1m×2m square regions S_1, S_2. We compared the recognition accuracy achieved by k-Nearest-Neighbor classifiers and Decision Tree in the confusion matrices shown in Table 3.10a and Table 3.10b respectively. The six states are detectable with an overall accuracy of 81.4% and 72.3% using kNN and Decision Tree classifiers respectively.

3.5 A Use Case: An Attention Monitoring System from FM Signal

3.5.1 Overview

Attention determines for a system the potential to impact the actions and decisions taken by an individual [99]. The management of attention covers the activation of attention as well as its detection and timely exploitation. The same action of the same system might be considered either as annoyance or be appreciated as helpful depending on whether the individual was focusing part or all of her attention towards the system or not. In the literature, we find various definitions that classify attention as well as its determining characteristics [94]. attention might be the tracking of gaze [100]. In general, aspects such as Saliency, Effort, Expectancy and Value are important indicators of attention [90]. this model and put a greater stress on the effort a person takes towards an object [22].

We consider the following scenario. In a corridor, a series of electronic poster frames are installed while people are walking by these frames. From the perspective of a specific poster, a significant part of its message shall be recognized by passers-by. Therefore, the poster should draw the attention of people passing by and, when this is achieved, it might possibly transport additional information. Consequently, the poster frame should be aware of people passing by, know where people are in order to attract attention at the right moment and detect whether attention is attracted. In this work we assume that the monitored individuals are not cooperating with the system and hence are not equipped with any part of the sensing hardware.

Such detection and management of attention may require elaborate installations and very specific sensors in order to accurately sense quantities such as Saliency, Effort, Expectancy and Value [99]. commercial installations, cost and ease of installation and not primarily the highest achievable accuracy are most important. Also more general, environmental sensors can provide sufficient information to estimate the attention state of individuals.

We propose to utilize ambient FM-radio signals for the detection of attention since it has a nearly perfect coverage in populated areas and the feature of cheap receiver hardware [53]. *But which aspects of attention can actually be captured by an fm-receiver?* Ferscha and others [22] discuss various aspects of attention and identify as most distinguishing factors changes in walking speed, direction or orientation. From FM-radio signals it is hard to detect the orientation of a person. However, it is feasible to classify walking speeds, walking direction or location of individuals. We show that with a straightforward installation, we can distinguish 1) an empty corridor, a person walking by and a person standing in front of a poster frame, 2) the specific poster the person is observing, 3) the location where a person is walking as well as 4) the walking speed of a person. We therefore argue that attention levels can be inferred upon interpretation of the changes in walking speed or direction as derived from our system. This information can provide some indication on the attention of persons towards a poster frame. Also, it can enable a frame to take action in order to catch the attention of a person just in the right moment.

3.5.2 System design

The design of attention monitoring system encompasses three main component as summarized in figure 3.8. We extract features for attention monitoring from FM-signals continuously broadcast by an FM-radio station.

The features are obtained from a series of continuous measurements $s_1, s_2, \ldots, s_t$ which are samples of the amplitude of ambient FM signals and grouped in windows $S_1, \ldots, S_n$ of k consecutive samples each $\langle s_{(i-1)k+1}, s_{(i-1)k+2}, \ldots, s_{ik} \rangle$. From these S_i, sets of features $\mathrm{F}_i = \langle f_{i,1}, f_{i,2}, \ldots, f_{i,m} \rangle$ are extracted and used for the monitoring of attention. The features we utilized are the mean (Avg_i), the variance (Var_i) and the energy (E_i) of S_i as detailed in Section 3.3.2.

These features have been derived among a greater set of features as well suited to achieve

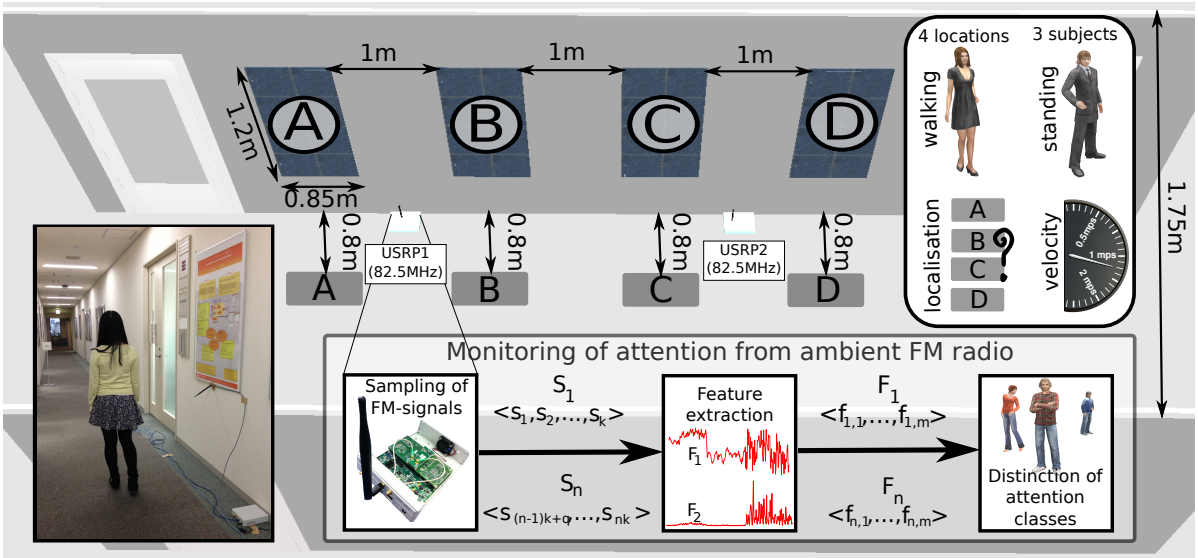


Figure 3.8: Sketch of the evaluation setting. The attention-monitoring system extracted features combining the data acquired by both USRP devices.

high accuracy for activity recognition conditioned on passive RF-based recognition systems in [75]. After the extraction of features, we randomly divided the collection F of all feature sets F_i into a training set Tr and a classification set Cl that met the conditions $\mathrm{Tr} \cup \mathrm{Cl} = \mathrm{F}$ and $\mathrm{Tr} \cap \mathrm{Cl} = \emptyset$. The set Tr is used to train the classifiers. After the training, classifiers will process Cl.

For a set of k activities $\mathcal{A} = \{a_1, \ldots, a_k\}$ let $\mathcal{I}(a_i)$, with $\mathcal{I}(a_1) \cup \mathcal{I}(a_2) \cdots \cup \mathcal{I}(a_k) = \mathrm{Cl}$, be the total number of instances for activity a_i and $\mathcal{I}_{\mathrm{cor}}(a_i)$ be the number of correctly classified instances for this activity in which the classification matches the ground truth. We define the accuracy by which an activity a_i can be detected as

$$\mathcal{ACC}(a_i) = \frac{\mathcal{I}_{\mathrm{cor}}(a_i)}{\mathcal{I}(a_i)}$$

3.5.3 Evaluation

In this section, we discuss case studies to demonstrate the viability of monitoring attention of people passing by several poster frames towards these frames. In all cases, we consider a

corridor with posters attached along one side (cf. figure 3.8).

Four posters of 0.85 m × 1.2 m which are separated by 1 m are attached alongside one wall of the corridor. We place the USRP devices between the two leftmost and rightmost poster frames on the floor. These N210[9] USRP devices are equipped with WBX[10] daughter boards and VERT900[11] antennas with 3 dBi antenna gain. Both devices continuously recorded the signal strength with a sample rate of 64 Hz, emitted by an ambient FM-radio station at 82.5 MHz while the attention of subjects towards the poster frames is monitored. We distinguish between four locations, 0.8 m in front of the posters (labelled A, B, C, and D) and the rest of the corridor. During the case studies, the subjects were walking along the corridor and through the marked areas or standing in front of one of the posters at the marked locations. As a baseline, the received signal from the empty corridor was recorded. For each action and all three subjects about two minutes of sample data each have been collected.

For all features, we utilize a window of 128 signal measurements, spanning a total of 2 seconds. Features are extracted from the data sets collected by USRP1 and USRP2 (cf. figure 3.8) and are merged for the distinction of attention classes. For this, we utilize a decision tree (DT) and a k-nearest-neighbour (k-NN) classifier from the Orange data mining Toolkit[12]. The k-NN classifier utilized 5 neighbors and weights their distance by the Euclidean distance. The decision tree utilises at minimum 10 instances in its leaves for pre-pruning and a recursive merge of leaves of the same major class with an m-estimate of 2 for post pruning. We apply a 10-fold cross validation.

3.5.3.1 State of the corridor

In our scenario, a corridor is equipped with electronic poster frames which shall detect the attention of passers-by and act accordingly. The most basic case to distinguish for the frames

[9] https://www.ettus.com/content/files/2987_Ettus_N200-210_DS_FINAL_1.27.12.pdf
[10] https://www.ettus.com/product/details/WBX
[11] https://www.ettus.com/product/details/VERT900
[12] http://orange.biolab.si/

Table 3.11: Mean accuracy for the distinction of the corridor states ‘empty’, ‘person standing’ and ‘person walking’

(a) Classification accuracy achieved by a k-NN classifier

		Classification		
		empty	standing	walking
Truth	empty	**.906**	.034	.06
	standing	.136	**.765**	.099
	walking	.021	.100	**.879**

(b) Classification accuracy achieved by a DT classifier

		Classification		
		empty	standing	walking
Truth	empty	**.877**	.064	.059
	standing	.041	**.852**	.107
	walking		.071	**.929**

is the state of the corridor. In particular, we consider whether the corridor is empty or occupied by a person and, when it is occupied, whether this person is walking or standing. In the case of electronic poster frames, the devices might change into an energy saving mode when the corridor is empty or also display more or less complex information conditioned on whether the person in the corridor is walking or standing.

Table 3.11 depicts the classification accuracy for these classes. For all classes, the mean classification accuracy over the sample windows of 2 seconds is near or above 0.8. In a second stage, we can now obtain information related to the attention of passers-by.

3.5.3.2 Focused attention towards specific frames

While walking by poster frames, brief snippets of the content can be grasped by individuals. However, an intense engagement with the more complex content of a poster requires a person to slow down her walking speed [22] and possibly come to a stand in front of the poster. We demonstrate the distinction in front of which poster a person is standing in the scenario depicted above. All parameters of the recognition system remain identical to section 3.5.3.1. All subjects have been standing and observing a poster at one of the locations labelled A, B, C or D in figure 3.8. The most characteristic feature to distinguish these cases is the mean of the signal strength. The average classification accuracy after 10-fold cross validation is depicted in table 3.12. We observe that the classification accuracy is in most cases above 0.9, in all

Table 3.12: Mean accuracy for the distinction in front of which poster a person is standing

(a) Classification accuracy achieved by a k-NN classifier

		Classification (Standing at)			
		Loc.A	Loc.B	Loc.C	Loc.D
Truth	Loc.A	**.876**	.011	.022	.090
	Loc.B		**.788**	.203	.008
	Loc.C		.063	**.929**	.009
	Loc.D	.009	.027		**.964**

(b) Classification accuracy achieved by a DT classifier

		Classification (standing at)			
		Loc.A	Loc.B	Loc.C	Loc.D
Truth	Loc.A	**.933**		.011	.056
	Loc.B		**.735**	.257	.009
	Loc.C		.080	**.911**	.009
	Loc.D	.054		.027	**.919**

cases it is near or above 0.8.

3.5.3.3 Tracking individuals in motion

While people are passing by poster frames in a corridor, a specific poster frame might have the intention to actively attract the attention of a passer-by. This attempt is most successful when the person is in the proximity of the poster, facing towards it. In order to optimally schedule such action, the location of the walking person has to be available at the system. We show that the location of a single person walking along a corridor can be traced by analyzing fluctuation of an incoming FM-radio signal. Similar to the case study detailed in section 3.5.3.2 we detect in front of which poster a person walking in the corridor is located.

Table 3.13 depicts our results. We observe that the classification of the location where a person is walking is harder than the classification of the location where a person is standing. However, the classification accuracy reached is still near or above 0.8.

Table 3.13: Mean accuracy for the distinction of walking at location A, B, C or D in the environment depicted in figure 3.8

(a) Classification accuracy achieved by a k-NN classifier

		Classification (Walking at)			
		Loc.A	Loc.B	Loc.C	Loc.D
Truth	Loc.A	**.779**	.118	.015	.088
	Loc.B	.107	**.804**	.071	.018
	Loc.C		.017	**.933**	.05
	Loc.D	.139			**.785**

(b) Classification accuracy achieved by a DT classifier

		Classification (Walking at)			
		Loc.A	Loc.B	Loc.C	Loc.D
Truth	Loc.A	**.754**	.228	.018	
	Loc.B	.175	**.772**	.035	.018
	Loc.C	.029	.017	**.953**	
	Loc.D	.125		.071	**.804**

3.5.3.4 Changes in the walking speed

As detailed in [22], an important indicator of the attention state of a person are changes in the walking speed. When a person is interested in a specific content of a poster, she might likely slow down to better perceive the content.

We obtain the walking speed of a passer-by from the fluctuation in ambient FM-radio signals. We collected for all three subjects and for three different velocities (0.5 m/s, 1 m/s, 2 m/s) samples of a duration of 2 minutes each. Again, k-NN and DT classifiers are utilized for training and classification. Table 3.14 illustrates our results. We observe that, although there is an indication towards the correct velocity in all cases, the accuracy greatly drops compared to the previous considerations. The confusion of these velocity levels especially for higher walking speeds is owing to the reduced duration an individual is located in front of a single poster during her walk. We can, however, achieve a higher recognition accuracy without increasing the distance between posters by abstracting from the 1 m/s walking speed

Table 3.14: Confusion matrices for the discrimination between walking speeds (0.5 m/s, 1 m/s, 2 m/s) achieved by k-NN and Decision Tree classifiers

(a) Classification accuracy achieved by a k-NN classifier

		Classification		
		0.5m/s	1m/s	2m/s
Truth	0.5m/s	**.641**	.219	.141
	1m/s	.109	**.453**	.438
	2m/s	.134	.284	**.582**

(b) Classification accuracy achieved by a DT classifier

		Classification		
		0.5m/s	1m/s	2m/s
Truth	0.5m/s	**.844**	.094	.063
	1m/s	.094	**.578**	.328
	2m/s	.164	.343	**.493**

Table 3.15: Confusion matrices for the discrimination between walking speeds (0.5 m/s, 2 m/s) achieved by k-NN and Decision Tree classifiers

(a) Classification accuracy achieved by a k-NN classifier

		Classification	
		0.5m/s	2m/s
Truth	0.5m/s	**.896**	.104
	2m/s	.219	**.782**

(b) Classification accuracy achieved by a DT classifier

		Classification	
		0.5m/s	2m/s
Truth	0.5m/s	**.925**	.075
	2m/s	.025	**.750**

(cf. table 3.15), distinguishing only between a slow walk and a running person.

Although we are then not able to distinguish the medium walking speed, note that the attraction of attention of a person in a hurry is not the intention of the considered system. Rather, we are focusing towards individuals in a relaxed, open state of mind to receive external stimuli and information.

Since the change in walking speed at a particular location might correspond to the attention level of passer-by, the information on the walking speed, monitored over time, can be utilized to grasp her attention level.

3.5.3.5 Altering the count of receive devices

In the above considerations, we have utilized two USRP devices since the experimental setting spans over five meters and the classification accuracy deteriorates with increasing distance to the receive antenna [66]. However, for economic reasons, a simple installation might be

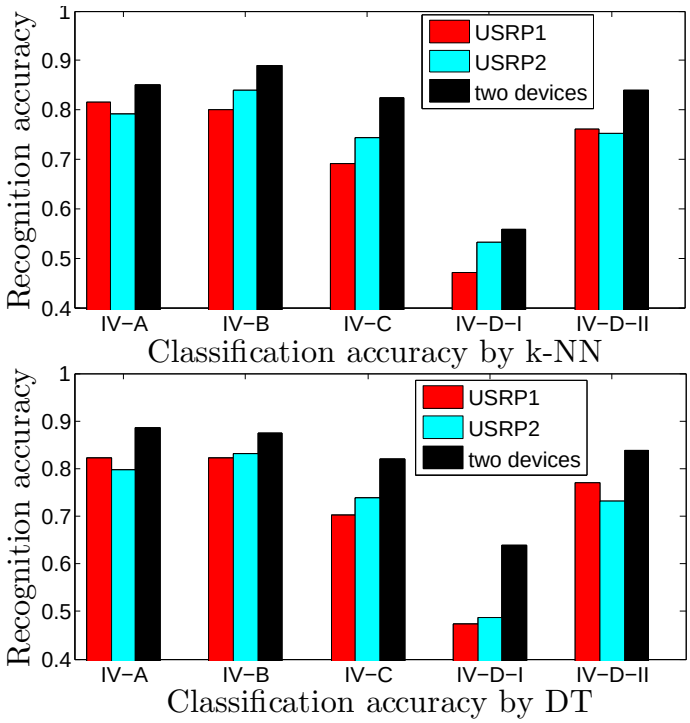


Figure 3.9: Comparison of the classification accuracy for the 4 cases described in section 3.5.3.1, section 3.5.3.2, section 3.5.3.3, and section 3.5.3.4 when only one receive device is utilized

designed in favour of only one receive device at the cost of a slightly reduced recognition accuracy for greater distances. In order to evaluate this impact for the monitoring of attention of passers-by within a corridor, we also consider the classification accuracy when the data from only one of the receive devices is utilized. The classification system and location of receive devices was not changed. Figure 3.9 depicts our results.

We observe that the classification accuracy benefits from the addition of the second device in all cases. With only one device, the overall classification accuracy drops by about 0.05 to 0.1 since the classification accuracy for individuals in greater distance deteriorates.

3.6 Summary

We have proposed a passive activity recognition method utilizing signal strengths originating from an ambient, non-controlled radio source, in this case a FM radio broadcasting station. To measure the performance of our method, several experiments conducted by 3 subjects each of which performed 5 activities of 'lying', 'standing', 'empty room', 'crawling' and 'walking' in prescript areas. In an $2m \times 2m$ area, exploiting two-stage method, which can remarkably improve the performance in comparison with one-stage method, overall accuracies of 85.6% and 84% are achievable for k-NN and Decision Tree classifier respectively. In this scenario, we also show that the performance of activity recognition will degrade with the increase of distance between the individual who performs activities and the transceiver. To validate the efficiency of our proposed activity recognition approach, we conducted another case study, which is a duplication of the previous experiments and the alternative scenario of which is schematized in fig 3.7. Experimental results demonstrated that for all classifiers comparable precision also can obtain in this scenario, indicating the universality of the proposed RF-based recognition. Furthermore, in this deployment 3.3, we explored whether the RF-based, passive, transceiver-free activity recognition method can reach about the same accuracies of distinction between four activities, specifically, standing, lying, walking and crawling as the generic approaches which leveraging accelerometer sensors for activity detection. In comparison with the detection means via accelerometer data, the experiments evaluated similar accuracies can also be obtained with our proposed method. Also, we discuss the device-free joint localization and activity recognition from FM radio. Upon the framework, we conduct a use case which enables to monitor the attention of a subject.

Chapter 4

Mining Location Contexts Using CSI information

4.1 Introduction

People resign a significant share of their time indoors (over 80%) [1], covering shopping in leisure time or attending meetings during office hours. Accurate information on mobility patterns and movement paths would enable improved building and path management and also help to advertise relevant information at right places. However, in most practical purposes, it can not be assumed that all subjects to be tracked can be equipped in advance, as this would require considerable resources. Recent research on RF-based indoor localization, however, has let to promising results. In these approaches, fluctuations in the ubiquitously available RF-Signals are exploited for localization and tracking purposes.

Traditional approaches require transceiver-equipped subjects and localize the device rather than its wearer by translating the observed changes and fluctuation in received signals to a coordinate system.

Examples are FM-based indoor localization [14], GSM-based techniques [47] as well as Bluetooth [20] or WiFi-based systems [4]. The reported localization accuracy of state-of-

art device-bound indoor systems is less than $0.5m$ [62], which meets the demand of most applications.

The main disadvantage of such approaches is, however, that all require a cooperating and equipped subject to be localized. However, in most practical purposes, it can not be assumed that all subjects to be tracked can be equipped in advance, since this would require considerable resources. A possible alternative is device-free passive (DfP) indoor localization [12, 97, 104]. In DfP indoor localization, fingerprint-based techniques are widely adopted, since the unpredictability of radio propagation due to multipath effects renders the alternative of analyzing RF signals challenging. A most recent work, *Pilot*, shows how correlation patterns of Channel-State-Information (CSI) from OFDM-MIMO links can be leveraged for accurate indoor localization [95]. Other state-of-the-art DfP localization systems include *Nuzzer* [60], and *PC-DfP* [98], both of which exploit the less accurate Received Signal Strength (RSS) indicator instead of the much finer grained CSI utilized in Pilot.

In this study, we further advance CSI-based indoor localization by adopting every single subchannel amplitude of CSI measurements and propose a single-stage direct classification. This is in contrast to Pilot, which utilizes the "*correlation*" between multiple CSI readings in a two-stage process. The contributions of this chapter are

a) a comprehensive analysis of the characteristics of CSI change induced by human presence.

b) a computationally efficient single-stage approach for indoor localization and tracking which takes full advantage of information provided by CSI measurements from commodity 802.11n WiFi wireless cards. The approach is computationally efficient and robust against high-dimensional CSI vectors due to PCA-based dimensionality reduction.

c) a large case study covering four diverse typical indoor spaces in two different buildings

d) a performance comparison to the state-of-art indoor positioning solutions *Nuzzer* [60], *PC-DfP* [98] and *Pilot* [95].

e) Kalman-filter and Particle-filter based approaches to track a moving target and a discussion on the adaptability of these two Bayesian filters to distinct walking patterns.

In particular, we have significantly revised and improved the analysis, conducted a more general experimental study in 4 diverse environments, added performance metrics for the analysis, compared the results achieved to two additional state-of-the-art algorithms, namely Nuzzer and PC-DfP and improved the overall discussion on the topic. Finally, and most significantly, our current system is able to continuously track a moving target. This is possible by the adaptation of Kalman-filter and Particle-Filter-based approaches. Both have been concisely investigated and are discussed in depth in this article.

The rest of this chapter is structured as follows. The preliminary studies are detailed in section 4.2. In section 4.3, we describe the implementation of our proposed CSI-based passive device-free indoor localization system. We detail the experiment setup in section 4.4. The evaluation results are presented in section 4.4.4. In section 4.5, we investigate the feasibility of tracking a moving person using Bayesian filtering. Finally, section 4.7 draws our conclusion and discusses future work.

4.2 Background and Motivation

4.2.1 Channel State Information

In the standard of the 802.11 protocol, RSS is defined as an indicator for the quality of a link, which characterizes the overall received signal power in the channel. With the wide adoption of multiple input multiple output-orthogonal frequency-division multiplexing (MIMO-OFDM) PHY technology in many wifi-class devices, RSS is no longer regarded as an accu-

rate metric, since the data streams are transmitted on various orthogonal subchannels independently and the quality of these subchannels differs one by one. In contrast to RSS, CSI contains link information in the granularity of a single MIMO-OFDM subcarrier. Therefore, it holds the potential for more accurate indoor localization. With the release of the CSI tool for commodity WiFi cards by Halperin et. al [27], we can aggregate both the amplitude and phase information for each MIMO-OFDM subcarrier. Let t, r be the number of transmit (TX) and receive (RX) antennas, and w the total number of subcarriers for a TX-RX pair. Based on the functionality of the CSI tool, a CSI vector can be obtained per packet, containing $t \cdot r \cdot w$ values of subchannels as

$$C = \{C_{i,k}^{m}\}, \quad i \in [1, t], \; k \in [1, r], m \in [1, w],$$

For each value $C_{i,k}^{m}$, it reflects both amplitude and phase of the RF signal $C_{i,k}^{m} = |C_{i,k}^{m}|e^{j \sin\theta}$ modulated at the subchannel m from transmit antenna i to receive antenna k.

4.2.2 Unpredictable Nature of CSI in Indoor Environments

It is of great importance to know the sources of errors and biases to design a robust indoor localization system. Regarding the RSS-based solutions, various experiments demonstrate that multiple effects are of concern for the location precision. For instance, in [97], Xu et. al have a subject stand within the Line-of-Sight (LOS) path between a single channel TX-RX link in a cluttered indoor environment and find that the blocking of the LOS path induced by humans may not lead to a decrease in RSS, which demonstrates the unpredictability of mulitipath fading induced bias. In the context of MIMO-OFDM, since the RSS is no longer a reliable indicator for the entire channel quality [27], in this section, we investigate the challenges posed by CSI measurements in solving the location estimate problem of cluttered environments.

To verify the multipath fading of subchannels, we carry out experiments in a typical do-

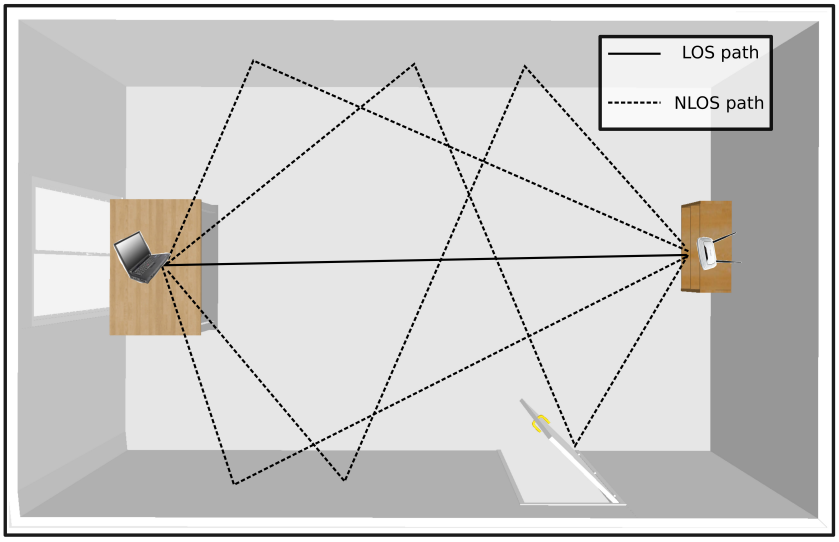


Figure 4.1: Experimental indoor environment

mestic home and install a transmitter (access point) and a receiver (laptop) within a distance of 4 meters at a height of 1.4m from the floor as depicted in Figure 4.1. As we need to explore the signal fluctuation of CSI measurements before and after blocking the LOS path, at first, the CSI measurements are recorded when no subject is present in the room, then we gather CSI readings while a subject stands in or out of the LOS path and extract the average values of non-subject occupied CSI measurements from them. Figure 4.2 illustrates the destructive and constructive probability of the CSI measurement changes from all 30 subchannels respectively when a subject blocks the LOS path in the area of interest. From this figure, we observe that the human-induced CSI change in LOS areas can be destructive as well as constructive. More specifically, for two certain subchannels, Figures 4.3a, and 4.3b show the histograms of the CSI amplitude change from subchannel 3 and 14 in both LOS and NLOS areas. From Figure 4.3a, we notice that the amplitude of CSI measurements in subchannel 3 decreases with probability of over 90% due to the blocking of the LOS path, while for subchannel 14, the strength of the signal attenuates with a probability smaller than 60%. From these figures, we conclude that the CSI measurements observed over diverse subchannels change in an unpredictable manner. Therefore, rather than employing a deterministic model of CSI mea-

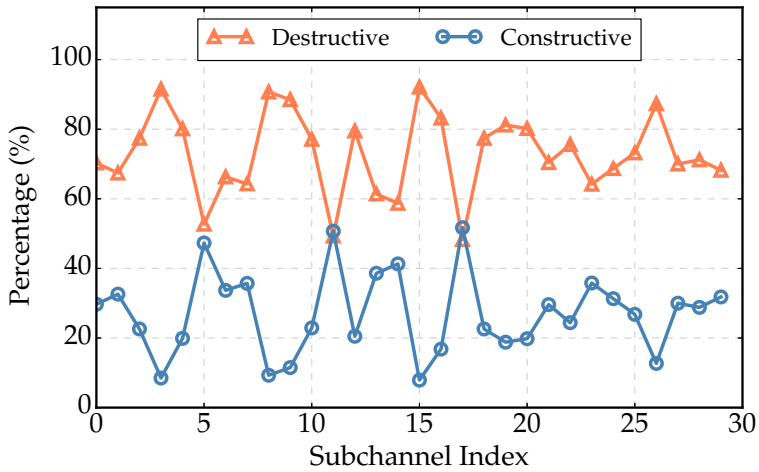


Figure 4.2: the probability from all the 30 subchannels that the CSI measurement changes result in constructive or destructive effect respectively when a subject blocks the LOS path in the area of interest

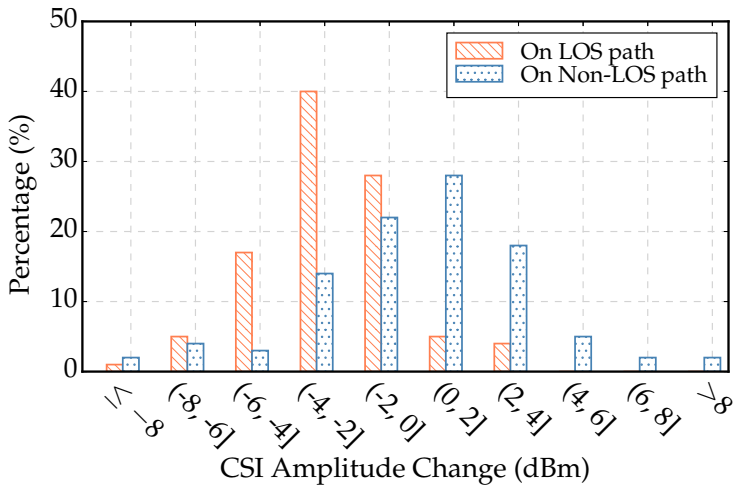
surements to estimate the distance, we propose to exploit a probabilistic approach for location discrimination.

4.3 Localization System Overview

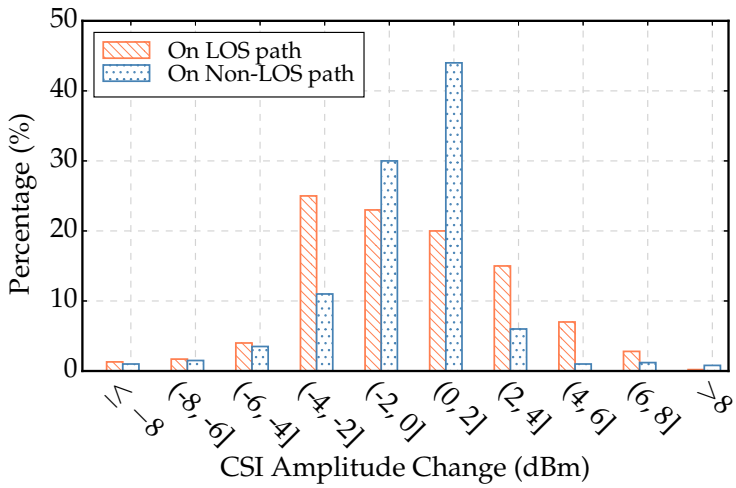
4.3.1 Problem Formulation

Here, we formulate the indoor localization problem and define the terms and variables used in our system. Let us consider a cluttered indoor environment, which is rich of multipath propagation. M transmitters and n receivers reside within the area of interest, forming a set $\mathcal{L}$ of $L = n \cdot m$ TX-RX links. We virtually divide the space into a set of small square cells with the same size, say a set $\mathcal{K}$ of K cells. For a link between a transmitter and a receiver $l \in \mathcal{L}$ and a square cell $k \in \mathcal{K}$, let $C_k(l)$ denote a CSI vector measured over link l when a subject is present at the area of cell k. There exist two phases for acquiring the CSI measurements as detailed below.

Training phase: During the training phase, a CSI measurement fingerprint map is built first.



(a) Measurements at Subchannel 3



(b) Measurements at Subchannel 14

Figure 4.3: Amplitude change of CSI readings in different subchannels

To construct it, we have a subject stand at cell k and collect a number of T CSI readings from all L links, forming a $L \times T$ matrix $C_k = \{C_k(1), ..., C_k(l), ..., C_k(L)\}'$, each element of which can be presented as a vector $C_k(l) = \{C_k^1(l), C_k^2(l), ..., C_k^T(l)\}$. After obtaining the CSI measurements at all cells $\mathcal{K}$, the fingerprint map consists of $\{C_1, C_2, ..., C_K\} = C_{\text{map}}$.

Test phase: During the online test operation, when a person is present in the monitored area, we gather a set of CSI measurements without the cell number, say C_{test} and determine

the cell where the person is standing given the information of C_{test} and C_{map}.

4.3.2 Probabilistic Method for Location Discrimination

As aforementioned, due to the random and unpredictable characteristics of CSI measurements, we propose a probabilistic fingerprint-based indoor localization system. More specific, we adopt Bayes estimator for the positioning problem. Since Bayes classifier is derived based on the Bayes' theorem, the conditional probability that a fingerprint $C_{\text{test}} = C$ collected during the test phase belongs to a certain cell $Y = y$ is given by

$$P(Y = y|C_{\text{test}} = C) = \frac{P(C_{\text{test}} = C|Y = y) \cdot P(Y = y)}{P(C_{\text{test}} = C)},$$

which is called posterior distribution. Obviously, the location discrimination problem of our system is to find the cell which maximizes the probability of posterior distribution, say $\hat{y} = \underset{y}{argmax}\, P(Y = y|C_{\text{test}} = C)$.

Since we suppose that the targeted person resides within the area of interest without any biased place, we can safely consider that each location y is equally likely, and thus the probability of $P(Y = y)$ is the same at every cell. Also, let $P(C_{\text{test}} = C)$ be identical for all possible locations y, as we suppose we gather fingerprints across all the cells with the same possibility during the test phase. To find the optimal solution for the posterior distribution $P(Y = y|C_{\text{test}} = C)$ is equivalent to having the maximum likelihood estimate of $P(C_{\text{test}}|Y)$. More specifically, the predicted cell for the location of the subject is therefore $\hat{y} = \underset{y}{argmax}\, P(C_{\text{test}} = C|Y = y)$.

As addressed in existing researches (cf. [15]), we model the amplitude of every CSI reading at location y to approximately follow a *multivariate Gaussian distribution* with mean μ_y and common covariance matrix Σ. Therefore likelihood estimate of $P(C_{\text{test}}|Y)$ is described

mathematically as

$$P(C_{\text{test}} = x|Y = y) = \frac{1}{\sqrt{(2\pi)^p|\Sigma|}} e^{-\frac{1}{2}(x-\mu_y)'\Sigma^{-1}(x-\mu_y)}, \tag{4.1}$$

where $p = L \cdot t \cdot r \cdot w$.

To estimate the parameters μ_y and Σ, we utilize the fingerprint map collected at training phase. The process is detailed as follows. At first, we convert Equation (4.1) to an equivalent description in the log-scale as

$$\delta(x|y) = -\frac{1}{2}\ln(|\Sigma|) - \frac{1}{2}(x-\mu_y)'\Sigma^{-1}(x-\mu_y) - \frac{y}{2}\ln(2\pi). \tag{4.2}$$

Given the set of T training CSI readings at cell y in the fingerprint map, $C_y = \{C_y^1, C_y^2, ..., C_y^T\}$, we assume each CSI measurement, C_y^t, $t \in [1, T]$, to be i.i.d[1]. Then taking the derivative w.r.t. μ_y of log-likelihood function (Equation (4.2)) and setting it to 0, we obtain

$$\hat{\mu}_y = \frac{1}{T}\sum_{t=1}^{T} C_y^t.$$

Similarly, we take its derivative w.r.t. Σ^{-1} of (Equation (4.2)) and let it be 0, leading to

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(C_y^t - \hat{\mu}_y)(C_y^t - \hat{\mu}_y)'.$$

If we substitute the two parameters $\hat{\mu}_y$ and $\hat{\Sigma}$ into Equation (4.1) and compute the probability at every cell, the cell in which we achieve the optimal probability, say $\hat{y}$, is exactly the estimated location of the subject.

[1] Our assumption is that every CSI reading has the same probability distribution and is independent with each other.

4.3.3 Dimensionality Reduction

To apply indoor location systems into some practical large-scale scenarios, one main concern that pose a challenge in positioning a subject in real time is the high dimensionality of data set. A large number of solutions have been proposed for the raised issue in RSS-based indoor localization systems (cf. [98], [16]). Since we leverage CSI measurements as the source signals, which has a much higher dimension than RSS indicators, so that it is a severer to solve the problem that the high dimensional CSI readings bring about.

In this study, we adopt *principal component analysis (PCA)* to project every CSI reading in the data set to a lower dimensional subspace. In particular, assuming that we would like to reduce each *p-dimensional* CSI measurement to a *q-dimensional* vector ($q < p$), we require to choose the q dimensionalities with largest variances and ignore the other less significant ones. These resulting *q-dimensional* features are called *principal components.*

For our system, we take the following procedures to calculate the q principal components. At each location y, we have T p-dimensional vectors which can be written as a $T \times p$ data matrix $\mathcal{D}$. Since PCA requires a mean-centered matrix $\mathcal{U}$ in order to calculate variations, the first step for projecting CSI vectors is to find the median vector γ of the data matrix $\mathcal{D}$ and subtract it from each row vector of $\mathcal{D}$ to obtain $\mathcal{U}$. Let $\mathcal{V}$ be the covariance matrix of $\mathcal{U}$, which can be computed by $\mathcal{V} = \mathcal{U}'\mathcal{U}/T$. Then we can calculate T eigenvectors and eigenvalues of $\mathcal{V}$. We can form the $T - q$ PCA projection matrix by choosing the q eigenvectors of $\mathcal{V}$ which have the q largest eigenvalues

$$\Phi_{pca} = [x_{1,}^{pca}\ x_{2,}^{pca} \dots x_{q}^{pca}].$$

For any p-dimensional CSI vector $\mathbf{d} = (d_1, d_2, \dots d_p)$, collected at location y, we project it into a corresponding q-dimensional vector by

$$\mathbf{d}_{pca} = (\mathbf{d} - \gamma) \cdot \Phi_{pca.}$$

4.4 Experimental Setup

4.4.1 Hardware Description and Data Aggregation

To evaluate the performance of our system, we deploy a wireless sensing networks to aggregate required CSI measurements as our testbed. In our study, we use TP-LINK WR841N wireless APs with an IEEE 802.11n compliant radio operating in the 2.4GHz unlicensed band as transmitters. The receivers utilized in our system are *lenovo* laptops integrated with Intel WiFi Wireless Link 5300 Cards. By using the modified driver released by [27], the receivers can probe one CSI reading per packet. The placement of APs and laptops is fixed and known a priori. Each laptop deployed in the targeted environment is synchronized with each other and receives 10 beacons from each AP per second.

In all the experiment settings, to aggregate the training data and construct the fingerprint map, we divide the indoor areas of interest into $0.75m \times 0.75m$ cells and record 100 CSI readings (approximately 10 seconds) from every TX-RX link in each cell. Note that all laptops can receive packets from all APs concurrently, since we run multiple processes to download data from the IP addresses of corresponding APs. To mitigate the bias induced by different orientation and movement, we collect 5 datasets (500 CSI measurements) in which the subject faces towards four different directions and walks randomly within a specific cell. During the test phase, a different subject is located at a random cell with random orientation. The subject may either stand still or walk at the cell. We collect CSI measurements at 500 locations (some locations are tested multiple times) and record samples with a duration of $5s$ (roughly 50 values) per cell.

4.4.2 Experiment Layouts

To confirm the validity of our system on various experimental scenarios, we carry out experiments in 4 different indoor environments. These environments are used for diverse functionalities and thus equipped with different domestic appliances and furnishings. For all four indoor spaces, transmitters and receivers are placed as depicted in Figure 4.4.

4.4.3 Performance Metrics

We introduce three performance metrics to evaluate the performance of our proposed device-free indoor localization system.

4.4.3.1 Cell Estimation Accuracy

The cell estimation accuracy quantifies the ratio of the number of cells where the positions of the target person are correctly estimated during the test phase to that of all cell locations and is calculated as

$$\epsilon_{\text{test}} = \sum_{i=1}^{N_{\text{test}}} I(y_i = \hat{y}_i)/N_{\text{test}},$$

where N_{test} is the total number of testing cell locations.

4.4.3.2 Median Distance Error

At the test phase, we consider the location of a cell to be misclassified if the estimated cell does not match with the human occupied cell. Under such circumstances, we take into account the average mismeasured distance between the center points of the estimated cells and actual ones, termed as median distance error. Formally, the median distance error is given by

$$\sigma_{\text{test}} = \sum_{i=1}^{N_{\text{test}}} ||y_i - \hat{y}_i||/N_{\text{test}},$$

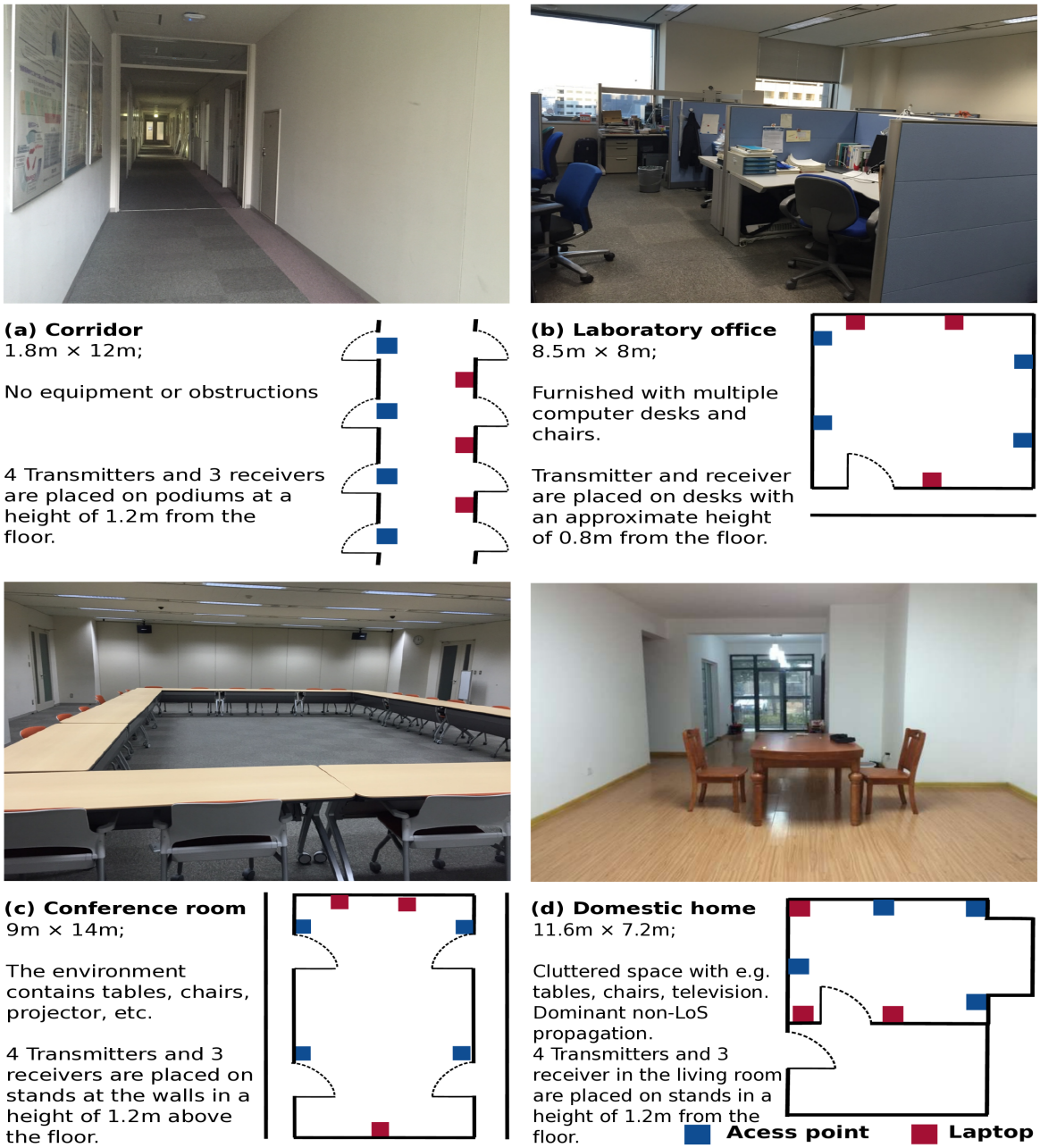


Figure 4.4: Layout sketches of our experimental indoor spaces. Environments (a), (b) and (c) are in the same building with walls mainly made of concrete. Environment (c) features a different penetration of walls which are made of wood and gypsum.

where $||y_i - \hat{y}_i||$ is the Euclidean distance between y_i and $\hat{y}_i$.

4.4.3.3 Average Processing Time

In our system, to address the problem of parameter estimation with high-dimensional datasets, we adopt PCA projection for dimension reduction. To validate the efficiency of PCA, the average processing time with different proportion of principle components to classify a testing location is measured on a MacBook Pro laptop (2.2 GHz Intel Core i7 processor, 16 GB 1600 MHz DDR3 memory, SSD storage).

4.4.4 Experimental Results

4.4.4.1 Comparing Various Localization Methods

In this section, we evaluate the performance of our system and compare it against other indoor localization systems including *Nuzzer* [60], *PC-DfP* [98] and *Pilot* [95]. The characteristics of these systems are summarized as follows:

Nuzzer: *Nuzzer* is a RSS-based device-free indoor localization system which constructs an offline radio map at the training phase and then estimates the location of an entity using a *Bayesian-based inference* algorithm. In *Nuzzer*, the distribution of RSS follows the Gaussian assumption.

PC-DfP: *PC-DfP* is also a Bayesian probabilistic classification based device-free indoor localization system using RSS measurements. To tailor the system to cluttered indoor environments, *PC-DfP* assumes that the density of the RSS mean vector of all the links at each location is *multivariate Gaussian* and adopts *linear discriminant analysis* for classification.

Pilot: *Pilot* is a two-stage device-free indoor localization which exploits the *correlation feature* of CSI measurements to achieve a better performance compared to RSS-based

schemes. *Pilot* will detect the presence of a human in the first stage and then trigger positioning phase to track the coordinate of the human in the second stage.

We conduct experiments in the 4 representative indoor environments as depicted in Section 4.4. Figure 4.5 and 4.6 illustrate the cell estimation accuracy and median distance error respectively. We should note that for *Pilot* it is possible that the system does not provide any prediction on the location. For this case, we define the error distance to be $min(width, length)$ of the testing room. In all indoor environments, we observe that our system outperforms the other three device-free localization systems. Compared to the RSS-based schemes (*Nuzzer*, *PC-DfP*), we attribute the better performance to more implicit information carried by CSI measurements in our system compared to RSS readings leveraged by *Nuzzer* and *PC-DfP*. Furthermore, in comparison with the other CSI-based localization system, *Pilot*, the performance gain of our system demonstrates that only one correlation feature and the two-stage location classification method adopted in *Pilot* are less effective than the location discrimination approach proposed in our system.

4.4.4.2 Impact of Principal Components

As aforementioned, parameter estimation is time-consuming due to the high-dimensional data of CSI vectors. In this section, we study the impact of using PLA dimensionality reduction on different evaluation metrics, including cell estimation accuracy (Figure 4.7), median distance error (Figure 4.8) and average processing time (Figure 4.9). From the three figures, we notice that a significant reduction in the PC-proportion (from 1 to $1/5$) barely worsens the the localization performance (see Figure 4.7 and 4.8), while dramatically saving processing time and thus improving the efficiency of our system (see Figure 4.9). Nonetheless, if we further reduce PCs (e.g. from $1/5$ to $1/10$), the classification accuracy will decline severely (see Figure 4.7) and error distance will hike sharply (see Figure 4.8); meanwhile the processing time can only reduce mildly, and its increasing rate progressively declines (see Figure 4.9). Thus, it is im-

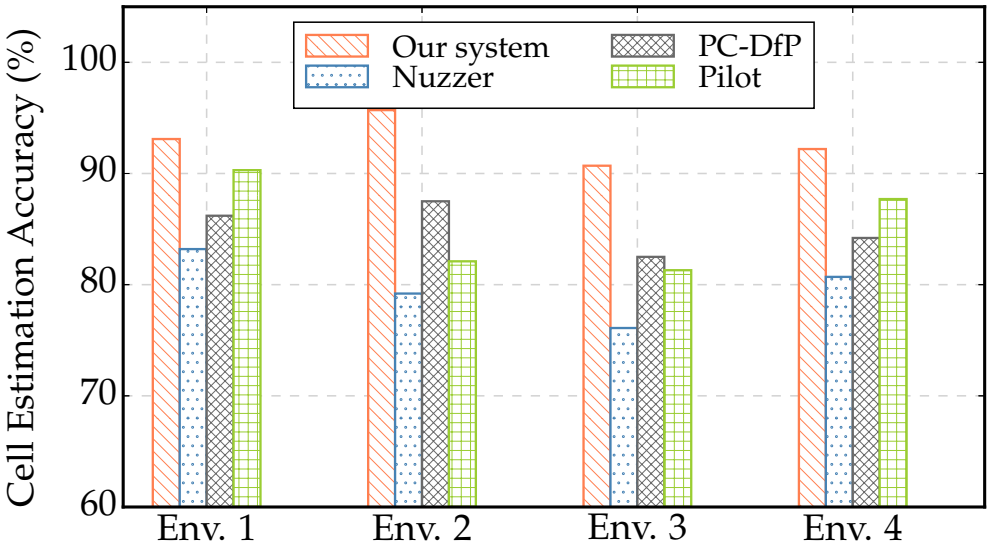


Figure 4.5: Comparison of cell estimation accuracy achieved by different indoor localization systems in four indoor environments

portant to choose an appropriate number of PCs to balance the accuracy-efficiency trade-off for indoor localization.

4.4.4.3 Influence of Data Aggregation Parameters

During the data aggregation procedure in our system, at each cell, we sample the CSI measurements from all the L TX-RX links, where each transmitter periodically broadcasts at a rate of P packets per second with a duration of T_{train} for the training phase and T_{test} for the test phase. Table 4.1 summarizes the default data aggregation parameters used in our experiments. In the following, we vary the values of these parameters and discuss the influence on the performance of our system.

4.4.4.4 Packet Reception Rate

From Figure 4.10a, we observe that we can achieve a cell estimation accuracy of over 90% by sampling the packets beyond a rate of 8 pks/s. This confirms a similar result reported in [73]

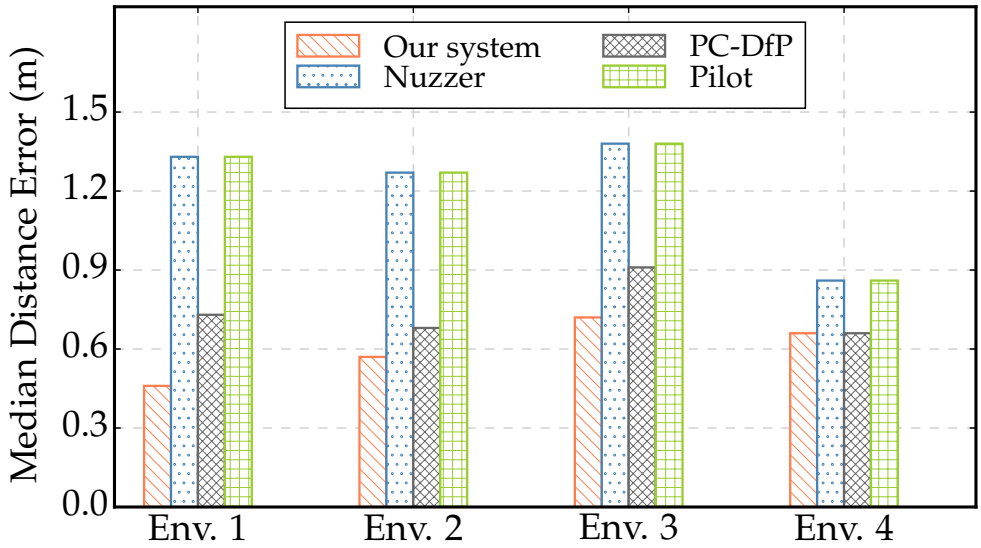


Figure 4.6: Comparison of median distance error achieved by different indoor localization systems in four indoor environments

for feasible RSS sample rates. Also, Figure 4.10b indicates that a reduced packet reception rate raises the performance of our system in terms of computation efficiency, which can be ascribed to the fewer data required to process a location.

4.4.4.5 Number of Links

Aggregating data from fewer TX-RX links can lead to significant reduction on the computational complexity of our system. However, it may also lead to a decrease in terms of localization accuracy. The experimental results achieved with different number of links are shown in Figure 4.11a and 4.11b. The results support the above conclusion that a smaller number of links deteriorates the cell estimation accuracy and reduces the localization processing time.

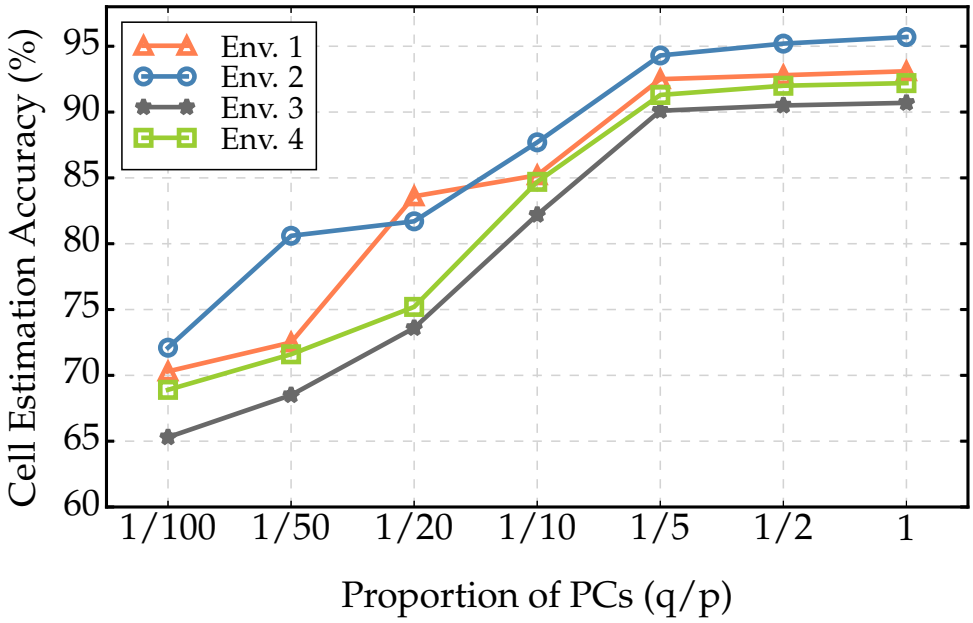


Figure 4.7: Cell estimation accuracy with different proportion of principle components (PCs) in four indoor environments

4.4.4.6 Time of Collecting Data

For fingerprint-based indoor localization schemes, the construction of a radio map is arduous, hence, reducing the time of collecting data at each cell can alleviate the effort during the training phase. With respect to the test phase, spending less time to collect the test measurements can make localization systems more responsive for locating the targeted people. While, localization precision may descend with a smaller number of gathered dataset.

Table 4.2 (Table 4.3) shows the average localization delay and the cell estimation accuracy/distance error with different time of collecting data during the training phase (test phase) using the same test dataset (training dataset) in all the indoor environments. There is a clear tradeoff between the cell estimation accuracy/distance error and localization delay with respect to the time of collecting data.

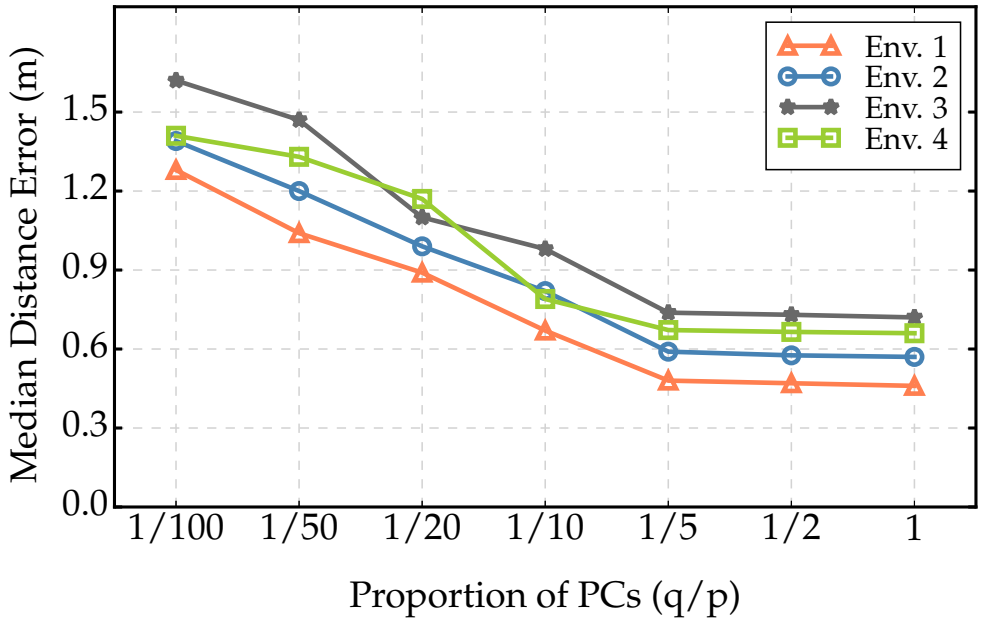


Figure 4.8: Median distance error with different proportion of principle components (PCs) in four indoor environments

4.5 Tracking With Bayesian Filtering

Using the coordinate of a person attained by the the positioning system we proposed as a basic, we continuously track the trajectory of the moving person leveraging two different Bayesian filtering techniques including Kalman filter and particle filter.

4.5.1 Kalman Filter Based Tracking System

Kalman filter is an optimal filter for linear dynamical systems, which maximizes the criterion of maximum a posteriori (MAP) for the recursive Bayesian filtering. To track a person in the environment of interest, we firstly model the problem and confirm that this model can be fitted into Kalman filtering conditions, then, for each t_h-second time frame, the two procedures, time update and measurement update, are performed iteratively.

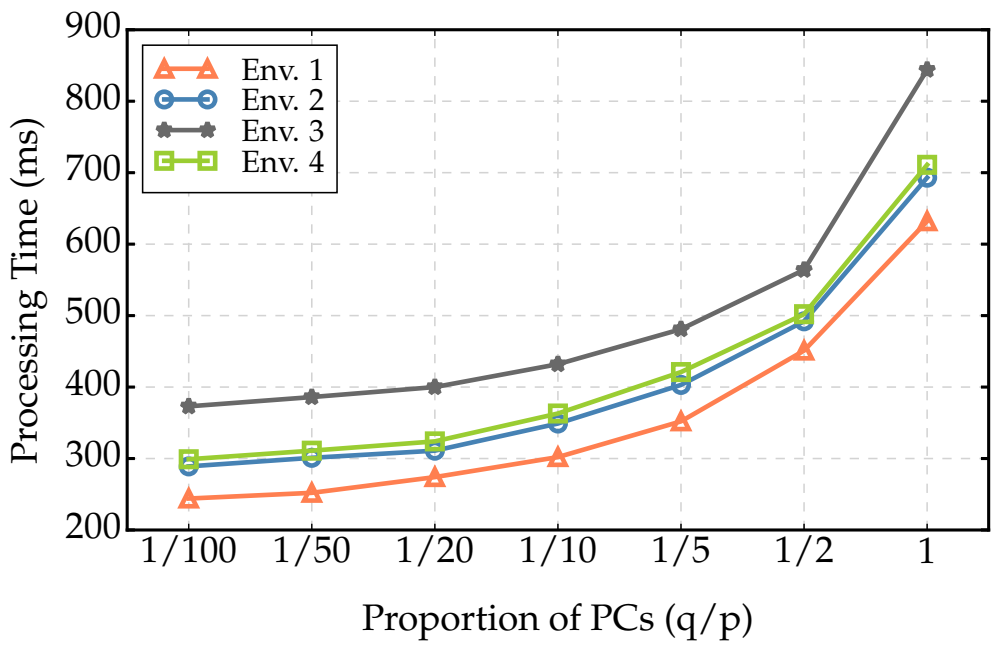


Figure 4.9: Average processing time to classify a test location with different proportion of principle components (PCs) in four indoor environments

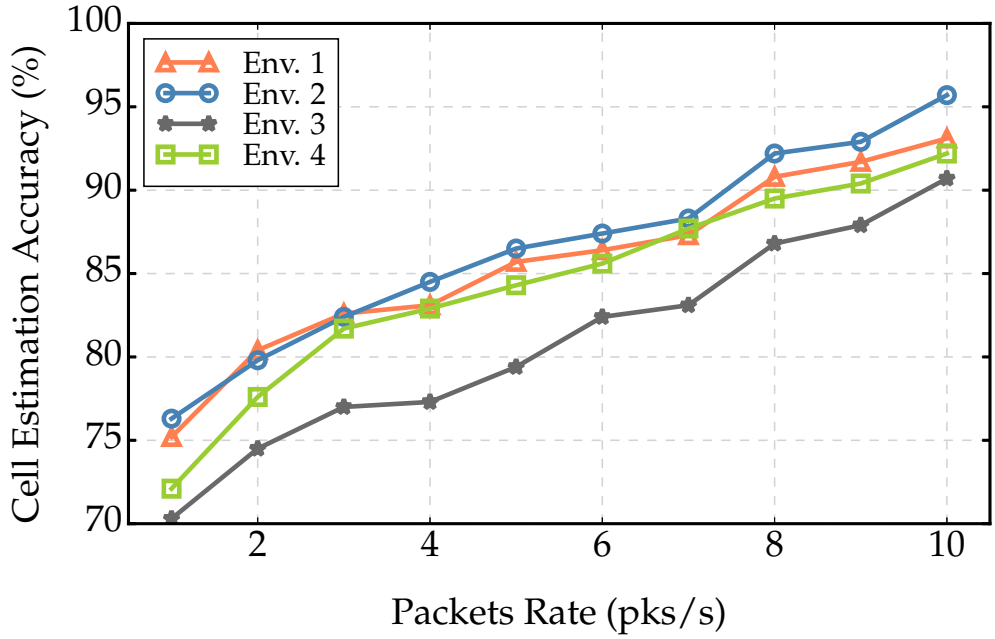
4.5.1.1 Modeling

The state equation of Kalman filter is

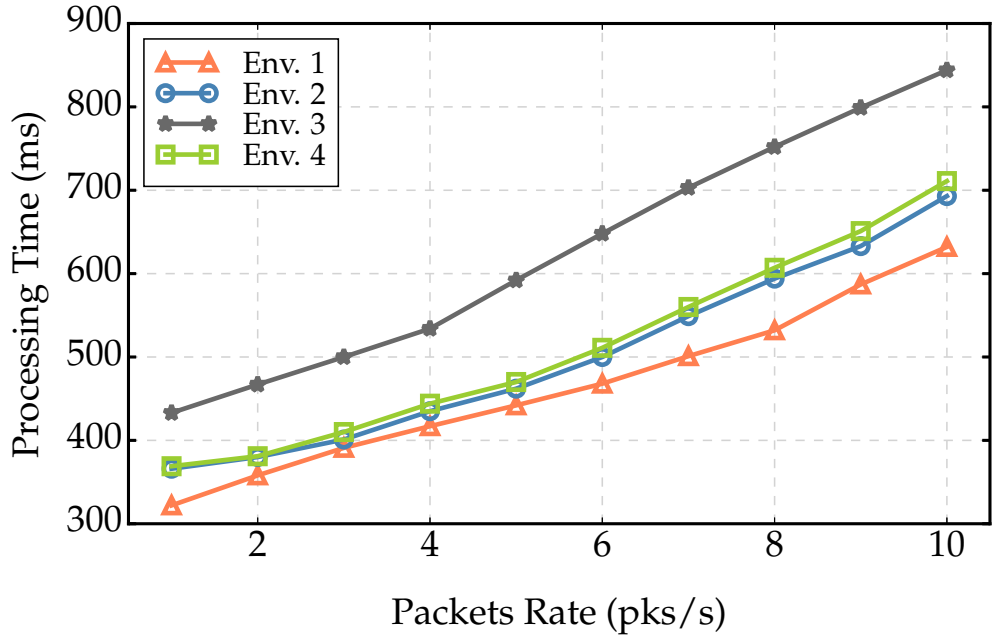
$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{w}_k,$$

where $\mathbf{x}_k$ is a *state vector* at the k-th iteration (at time instant kt_h), $\mathbf{A}$ is the *state transition matrix* and $\mathbf{w}_k$ is the process noise. For our tracking system, we include both coordinate (px_k, py_k) and velocity (vx_k, vy_k) in the state vector $\mathbf{x}_k = [px_k, py_k, vx_k, vy_k]'$. The state transition matrix can be presented as

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

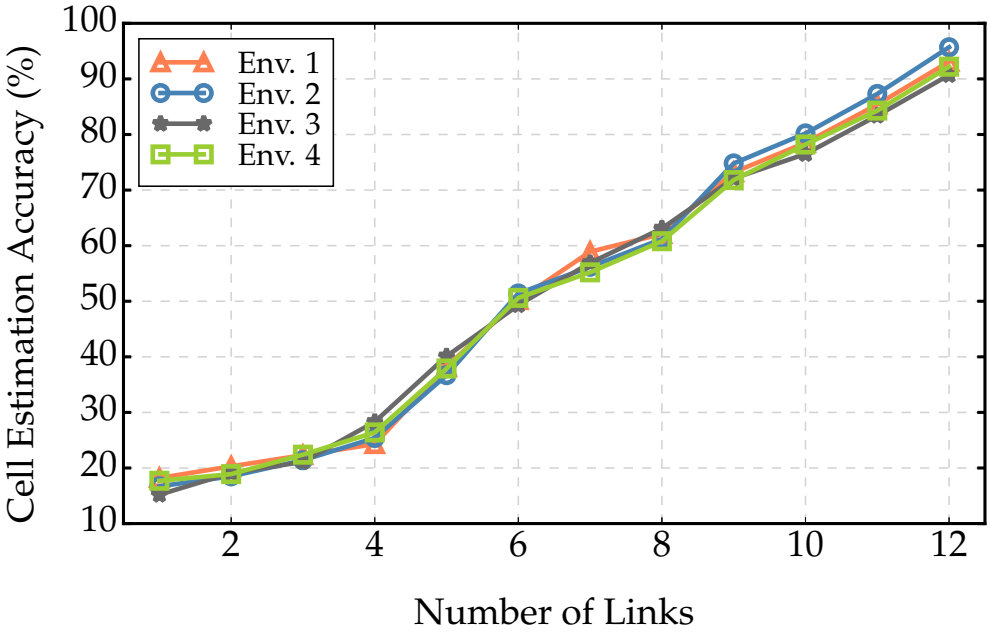


(a) Cell estimation accuracy with different packet reception rate

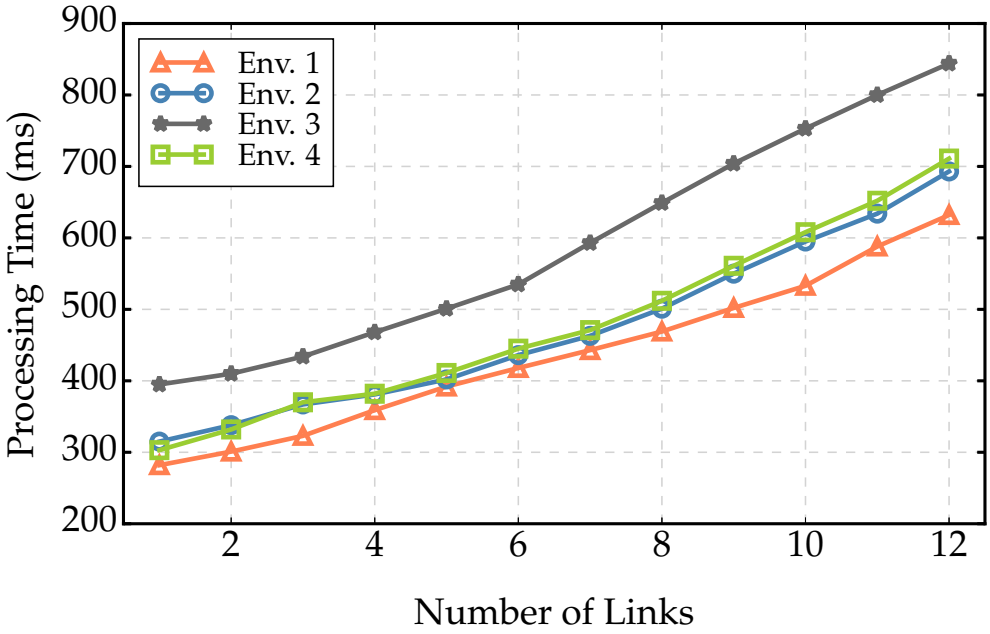


(b) Average processing time to classify a test location with different packet reception rate

Figure 4.10: Influence of packet reception rate on the performance of our indoor localization system



(a) Cell estimation accuracy with different number of links



(b) Average processing time to classify a test location with different number of links

Figure 4.11: Influence of number of links on the performance of our indoor localization system

Table 4.1: Default data aggregation parameters used in our experiments

Parameter	**Notation**	**Default value**
Number of links between TX-RX pairs	L	12
Packet broadcasting rate per second	P	10
Time duration for data aggregation per cell at training phase	T_{train}	$10s$
Time duration for data aggregation per cell at test phase	T_{test}	$5s$

Table 4.2: Influence of time frame of data aggregation on the performance of our indoor localization system during training phase

T_{train} (s)	**2**	**4**	**6**	**8**	**10**
Cell estimation accuracy $(\%)$	84.4	88.5	88.9	91.7	92.9
Distance error (m)	0.77	0.72	0.73	0.68	0.60
Processing time (ms)	443	491	507	656	720

and $\mathbf{w}_k = [0, 0, wx_k, wy_k]'$.

Similarly, the measurement equation is formulated as

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{u}_k,$$

where $\mathbf{z}_k$ is the measured coordinate of the tracked person from our localization system at the k-th iteration, $\mathbf{u}_k = [ux_k, uy_k]'$ indicates the measurement noise and the measurement matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Table 4.3: Influence of time frame of data aggregation on the performance of our indoor localization system during test phase

T_{test} (s)	**1**	**2**	**3**	**4**	**5**
Cell estimation accuracy $(\%)$	79.5	82.1	89.8	91.7	92.9
Distance error (m)	0.86	0.80	0.74	0.66	0.60
Processing time (ms)	473	534	590	655	720

For Kalman filter, both $\mathbf{w}_k$ and $\mathbf{u}_k$ follow zero-mean normal distributions with covariance $\mathbf{Q}_k$ and $\mathbf{R}_k$ respectively.

4.5.1.2 Update

In each iteration, we start a two-step process: *time update* and *measurement update*. In the time update process, state vector and error covariance are predicted according to estimated values of the previous iteration. Then, in the measurement update process, the estimated state vector and error covariance of the current iteration are corrected using the new measurement.

a) *Time Update*: in our tracking system, the two prediction equations for Kalman filter are given by

$$\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}_{k-1}, \tag{4.3}$$

$$\mathbf{P}_k^- = \mathbf{P}_{k-1} + \mathbf{Q}_k. \tag{4.4}$$

In Equation (4.3), $\hat{\mathbf{x}}_{k-1}$ is the *posterior estimate* of state vector $\mathbf{x}$ after the $(k-1)$-th iteration which means $\hat{\mathbf{x}}_{k-1}$ has been corrected at the measurement update procedure. $\hat{\mathbf{x}}_k^-$ denotes the *prior estimate* which is the prediction of a state vector at time kt_h without measurement correction. In State Covariance Prediction Equation (Equation (4.4)), the *prior error covariance* $\mathbf{P}_k^-$ is updated based on the covariance of process noise $\mathbf{Q}_k$ and

posterior error covariance of $(k-1)$-th iteration $\mathbf{P}_{k-1}$.

b) *Measurement Update*: the measurement update process for Kalman filter can be described by the following step:

Innovation Covariance: $\mathbf{S}_k = \mathbf{H}\mathbf{P}_k^{-}\mathbf{H}' + \mathbf{R}_k$,

Kalman Gain: $\mathbf{K}_k = \mathbf{P}_k^{-}\mathbf{H}'\mathbf{S}_k^{-1}$,

Measurement Residual: $\hat{\mathbf{u}}_k = \mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^{-}$,

State Estimate Update: $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^{-} + \mathbf{K}_k\hat{\mathbf{u}}_k$,

State Estimate Covariance: $\mathbf{P}_k = (I - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^{-}$.

4.5.2 Particle Filter Based Tracking System

In practice, it is of crucial importance to consider the elements of nonlinearity and non-Gaussian distribution to model the tracking systems accurately. In this study, we consider particle filters as the approximate nonlinear Bayesian filtering. In particular, since the conventional sequential importance sampling (SIS) based method for particle filtering often has the *degeneracy problem*, we track the targeted person using sequential sampling importance resampling (SIR) filter, which is a state-of-art robust particle filtering method and widely applied to recursive Bayesian filtering problems [41]. Compared to Kalman filter, the assumptions required to formulate recursive Bayesian filtering problems into a SIR filter are much weaker. We only need to know the state dynamics distribution $p(\mathbf{x}_k^m \mid \mathbf{x}_{k-1}^m)$, likelihood function $g(\mathbf{y}_k \mid \mathbf{x}_k)$[2] and select an appropriate importance density function $q(\mathbf{x}_k \mid \mathbf{x}_{k-1}^m, \mathbf{y}_k)$ and initial state estimates. Given the aforementioned assumptions, the SIR filter leveraged in our tracking system is described by Algorithm 4.1.

[2] $g(\mathbf{y}_k \mid \mathbf{x}_k)$ is equivalent to the measurement function $\mathbf{y}_k = h(\mathbf{x}_k) + v_k$. Since the measurement function can often be modeled, it is reasonable to assume $g(\mathbf{y}_k \mid \mathbf{x}_k^m)$ to be known.

Algorithm 4.1 Sequential Importance Sampling with Resampling

• Draw a particle ξ_0^m from $\xi_0^m \sim p_0(\mathbf{x}_0)$ and set $w_0^m = 1/M,\ m = 1, \ldots, M$
▷In the first place, initialize M particles from the distribution $p_0(\mathbf{x}_0)$ and set their weight is equally likely.
• **For** $k = 1, \ldots, N$ recursively **do**

1. Sampling
– Draw $(\xi_k^m,\ m = 1, \ldots, M)$ from $\xi_k^m \sim p(\mathbf{x}_k \mid \xi_{k-1}^m)$
2. Weight Calculation
– Compute the updated importance weights

$$w_k^m = \frac{g(\mathbf{y}_k \mid \xi_k^m) p(\xi_k^m \mid \xi_{k-1}^m)}{q(\xi_k^m \mid \xi_{k-1}^m, \mathbf{y}_k)}$$

3. Resampling
– Normalize importance weights

$$w_k^m = \frac{w_k^m}{\sum_{n=1}^{M} w_k^n},\ m = 1, \ldots, M$$

– Resample particles according to importance weights

$$\{\xi_k^m,\ 1/M\} \leftarrow \{\xi_k^m,\ w_k^m\},\ m = 1, \ldots, M$$

• **End for**

In our tracking system, we suppose the density of initial particles to be zero-mean Gaussian with variance σ^2, $p_0(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0;\ 0, \sigma)$. Therefore, in the initialization phase, we generate M particles, ξ_0^m, $(m = 1, \ldots, M)$, each of which, is i.i.d drawn from $\mathcal{N}(0, \sigma)$. We also assume the weight of each particle is uniformly distributed, so that $w_0^m = 1/M$. As demonstrated in Algorithm 4.1, three procedures are iterated for each t_h-second time interval: 1) *sampling*, 2) *weight calculation* and 3) *resampling*. In the sampling process, given the coordinate of the m-th particle at the $(k-1)$-th iteration ξ_{k-1}^m, the state dynamics distribution of our system $p(\mathbf{x}_k \mid \xi_{k-1}^m)$ assumes to be a bivariate Gaussian distribution

$$p(\mathbf{x}_k \mid \xi_{k-1}^m) = \mathcal{N}(\mathbf{x}_k;\ \xi_{k-1}^m + v_e(\xi_{k-1}^m) \cdot t_h,\ \textstyle\sum_{k-1}^m),$$

where $v_e(\xi_{k-1}^m)$ is the velocity of the m-th particle at the $(k-1)$-th iteration and $\sum_{k-1}^m$ is the

covariance matrix of this distribution. We approximate the velocity as $v_e(\xi_{k-1}^m) = \xi_{k-1}^m - \xi_{k-2}^m$. By using this estimation method, our SIR tracking system is feasible in the situation where the state dynamics function is not linear, so that the walking speed of the target person changes at different time intervals. In the weight calculation process, the original weight update equation of the SIR filter is shown to be

$$w_k^m \propto w_{k-1}^m \frac{g(\mathbf{y}_k \mid \xi_k^m) p(\xi_k^m \mid \xi_{k-1}^m)}{q(\xi_k^m \mid \xi_{k-1}^m, \mathbf{y}_k)}.$$

Therefore, the weight update equation can be given by

$$w_k^m \propto g(\mathbf{y}_k \mid \xi_k^m).$$

In our system, we model the measurement function as

$$\mathbf{y}_k = \mathbf{x}_k + v_k,$$

where $v_k \sim \mathcal{N}(0, \varphi)$, hence, $w_k^m = \mathcal{N}(y_k^m; \xi_k^m, \varphi)$. In the resampling phase, to mitigate the degeneracy problem induced by the particles with small weights, we resample M particles according to their weights and ensure that the new weights of the resulting particles are distributed evenly with probability $1/M$. We implement the resampling algorithm based on order statistics [10].

4.6 tracking results

4.6.1 Data Set

In the set of experiments, we collected datasets of CSI measurements when the targeted person is moving in all four indoor environments as depicted in Figure 4.4. More specifically, at each

environment, the subject walks along a random trajectory in the spaces following two walking types: (1) fixed velocity (approximately $2m/s$) and (2) free-style velocity, at which the subject can walk at any speed or stand at a certain location, for about 5 minutes. To obtain the ground-truth information where the subject is present, we use a smartphone to record the videos. Table 4.4 lists the experimental parameters used in the study.

Table 4.4: Experimental parameters used in our tracking system

Parameter	Notation	Value
covariance of process noise for Kalman filter	$\mathbf{Q}_k$	$2I_4$
covariance of measurement noise for Kalman filter	$\mathbf{R}_k$	$5I_2$
number of particles for particle filter	M	1000
time frame of each iteration for both filters	t_h	$1s$

4.6.2 Tracking Accuracy

We use both filters described above to track the trajectories of the subject with two different walking styles. Table 4.5 shows the mean absolute errors (MAEs) of various methods for each environment. As illustrated in this table, among all situations, applying the Kalman filter to the dataset of the subject moving with fixed velocity, achieves the best tracking precision with an average MAE of $0.63m$. That is a $0.17m$ improvement compared to particle filter based tracking with the same dataset. Whereas, when the subject mimics the walking style in real life (walk at random speed and stop freely), the average MAE at all the four indoor environments using particle filter is $0.92m$, a $0.31m$ improvement compared to that using Kalman filter. These results validate that Kalman filter is an optimal solution for linear dynamics systems (constant walking speed), in which all noise satisfies the normal distribution (we assume $\mathbf{Q}_k$ and $\mathbf{R}_k$ are Gaussian noise). Under nonlinear circumstances (free walking), the experimental

Table 4.5: Average MAEs (in meters) from our proposed tracking approaches

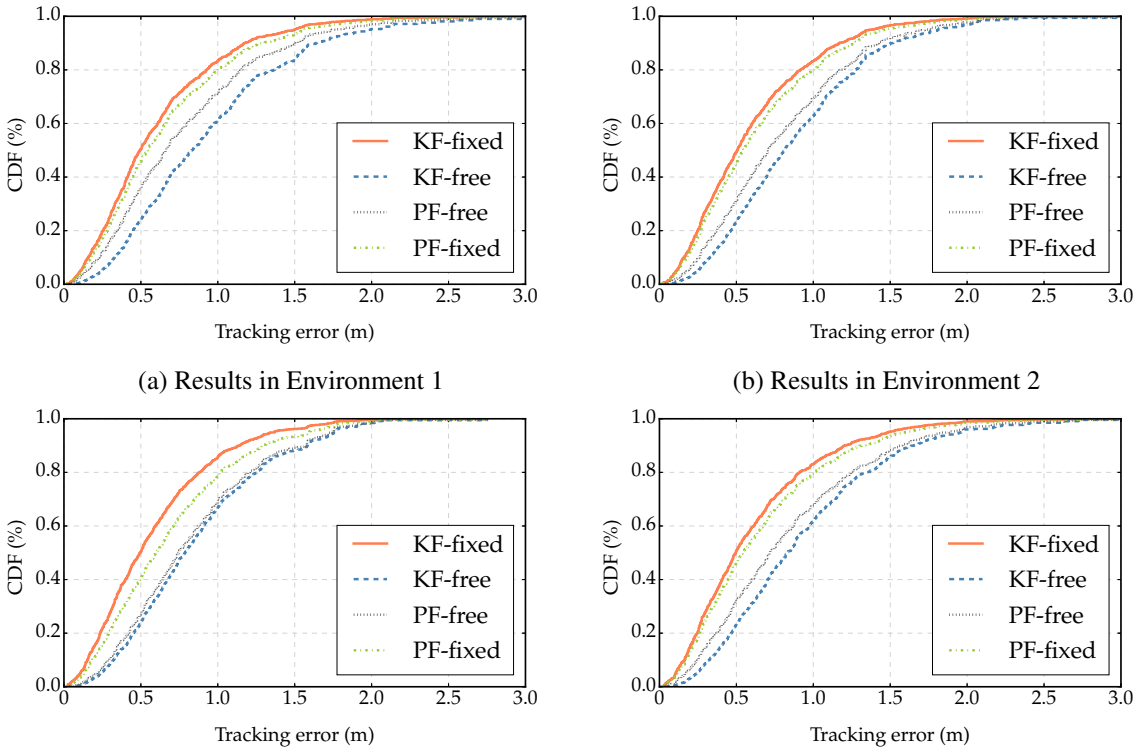
	Env.1	**Env.2**	**Env.3**	**Env.4**
Kalman filter (fixed velocity)	0.57	0.66	0.77	0.63
Kalman filter (free-style velocity)	0.91	0.97	1.13	1.02
particle filter (fixed velocity)	0.63	0.70	0.81	0.69
particle filter (free-style velocity)	0.86	0.88	0.94	0.91

results support the conclusion that particle filter is to be preferable to Kalman filter.

The cumulative distribution functions (CDFs) of the tracking error at each environment are shown in Figure 4.12. All CDFs also show that the Kalman filter achieves better performance than the particle filter when the subject is moving with immutable speed; however, particle filter outperforms Kalman filter once the walking pattern of the person is unpredictable. The consistent experimental results in the four distinct indoor environments also imply that our conclusion can be ubiquitously applied to indoor spaces.

4.7 Summary

We have presented a fingerprint-based device-free system that enables precise localization in indoor spaces. In our system, we aggregate the CSI measurements from commodity 802.11n WiFi devices, so that fine-grained subchannel information can be utilized to localize a subject. Classification is done comparing testing CSI readings with the CSI fingerprints and determine the location with highest probability by Bayes Classification. The performance of our system can be further enhanced by reducing dimensionality with PCA. The experimental evaluation in four different indoor environments shows that the system can outperform the state-of-the-art systems including *Nuzzer*, *PC-DfP* and *Pilot* in terms of both cell estimation accuracy and



(a) Results in Environment 1

(b) Results in Environment 2

(c) Results in Environment 3

(d) Results in Environment 4

Figure 4.12: The cumulative distribution functions (CDFs) of tracking errors of our proposed tracking approaches in different indoor environments

error distance. We further apply Kalman filter and Bayesian filter for tracking the moving target and show that Kalman filter is a preferred option for tracking the subject with fixed walking speed, while Bayesian filter is more robust to these scenarios where the subject walks randomly.

An intriguing open future direction is the localization and tracking of multiple persons simultaneously. Also, we are working on making the system resilient to changing indoor environments.

Chapter 5

Conclusion and Future Work

In this chapter, I discuss the limitations of the proposed solutions in this dissertation, summarize the several contributions of this thesis and present the paths for future research.

5.1 Discussion

Understanding the location and activity of a subject is an important requirement for context-aware systems. With the upcoming era of "internet of things", there will be increasing number of devices equipped with assorted sensors, which can gather information about the environment and facilitate a larger range of context-aware applications. Virtually, all IoT devices have the radio frequency (RF) interface to access the spectrum which provides us a new sensing modality context-aware applications. In this thesis, we have succeeded in recognizing the simple activities of a subject and positioning an individual by analyzing the various characteristics of RF signals interfered with different activities or at different locations. While my proposed two systems take a step in this direction, much more work remains to be done. For example, both systems require a training phase. However, some error can be introduced into sensed data which can result in the uselessness of training data. Therefore, the calibration is an issue we have not addressed yet. Also, in the approaching of IoT, there must be much

more devices anywhere in our ambient living environments. However, since the spectrum is scarce, devices might operate at the same spectrum band, therefore, our proposed systems should evolve to tackle with the channel interference problem. Even worse, a malicious jammer may intentionally disrupt the ability of a receiver to receive information that is intended to be transmitted to that receiver. In this case, the intentional radio interference caused by jamming attack can obviously threat the security of localization system, which pose another issue to the design of our system.

5.2 Conclusion

In the research area of context-aware computing, my thesis makes several contributions as detailed in the following:

- We confirm the interference of human's activities with FM-radio signals can lead to the different patterns of variations of RF signals. Based on these fundamental observations, we propose a passive FM-signal based activity recognition scheme. Following the scheme, we succeed in the detection of 5 activities using classification algorithms including k-NN and decision tree. We also experimentally demonstrate our system can simultaneously locate a subject and detect the activity information of this subject. Based upon the aforementioned framework, we design an attention-monitoring system, which enables classification of different characteristic patterns on amplitude-based features induced by human's locomotion. Our experiments demonstrate that for the distinction between an empty corridor or a corridor whether there is a subject facing a poster display, our system can acquire an accuracy of over 85%. Also, we show that our algorithm can discriminate several levels of walking velocities.

- The research on indoor localization has received great interest in recent years. This has been fuelled by the ubiquitous distribution of electronic devices equipped with a radio

frequency (RF) interface. Analyzing the signal fluctuation on the RF-interface can, for instance, solve the still open issue of ubiquitous reliable indoor localization and tracking. Device bound and device free approaches with remarkable accuracy have been reported recently. In this thesis, we present an accurate device-free passive (DfP) indoor location tracking system which adopts Channel state Information (CSI) readings from off-the-shelf WiFi 802.11n wireless cards. The fine-grained subchannel measurements for MIMO-OFDM PHY layer parameters are exploited to improve localization and tracking accuracy. To enable precise positioning in the presence of heavy multipath effects in cluttered indoor scenarios, we experimentally validate the unpredictability of CSI measurements and suggest a probabilistic fingerprint-based technique as an accurate solution. Our scheme further boosts the localization efficiency by using principal component analysis (PCA) to filter the most relevant feature vectors. Furthermore, with Bayesian filtering, we continuously track the trajectory of a moving subject. We have evaluated the performance of our system in four indoor environments and compared it with state-of-art indoor localization schemes. Our experimental results demonstrate that this complex channel information enables more accurate localization of non-equipped individuals.

5.3 Future Work

In this section, I discuss next steps for this work. First, I think our passive FM-based detection system can be capable of recognizing the activities of multiple subjects, but which I have not yet validated. Second, the up-to-date CSI measurement tools [85, 96]enable stitching multiple spectrum bands and extracting a much larger number of CSI measurements from subchannels. Therefore, I think the accuracy of our proposed CSI-based indoor localization system can further enhance with these new CSI measurement tools.

The specific possible directions of my work are as follows:

- One advantage of device-free passive detection systems is that these systems are able to localize multiple people at the same time, since all the subjects interfere with RF signals and induce various changing patterns of radio frequency. While my proposed framework only considers to detect the activities of a subject, I believe recognizing the activities of multiple subjects could be completed via our system by designing appropriate algorithms. One key problem to tackle with is nonlinear fading effects of multiple subjects. Also, as presented in the studies [32, 50], RSS measurements on the links between commercial wireless devices are utilized estimate their breathing rate, future direction of our thesis can detect the breath rate using other RF sources.

- In this thesis, the CSI measurement tool we utilize only support the functionality of extracting the CSI measurements of one particular spectrum band, however, according some latest systems [85, 96], it is possible to splice the CSI measurements from multiple WiFi frequency bands. Upon these systems, much more CSI measurements can be gained for indoor localization as raw inputs. Therefore, we believe performance of our CSI-based indoor system can be improved with these up-to-date CSI measurement tools.

References

[1] Source: Strategy Analytics: http://www.strategyanalytics.com.

[2] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles. Towards a better understanding of context and context-awareness. In *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing*, HUC '99, pages 304–307, London, UK, UK, 1999. Springer-Verlag.

[3] F. Adib and D. Katabi. See through walls with wifi! In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*, SIGCOMM '13, pages 75–86, New York, NY, USA, 2013. ACM.

[4] P. Bahl and V. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784 vol.2, 2000.

[5] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *Proceedings of PERVASIVE 2004*, volume LNCS 3001, 2004.

[6] M. Berchtold, M. Budde, D. Gordon, H. R. Schmidtke, and M. Beigl. Actiserv: Activity recognition service for mobile phones. In *International Symposium on Wearable Computers (ISWC)*, 2010.

[7] M. Berchtold, M. Budde, H. R. Schmidtke, and M. Beigl. An extensible modular recognition concept that makes activity recognition practical. In *Proceedings of the 33rd annual German conference on Advances in artificial intelligence*, KI'10, pages 400–409, Berlin, Heidelberg, 2010. Springer-Verlag.

[8] 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.

[9] A. T. Campbell, E. Larson, G. Cohn, J. Froehlich, R. Alcaide, and S. N. Patel. Wattr: A method for self-prowered wireless sensing of water activity in the home. In *Proceedings of the 12th international conference on Ubiquitous computing (UbiComp 2010)*, 2010.

[10] J. Carpenter, P. Clifford, and P. Fearnhead. Improved particle filter for nonlinear problems. *Radar, Sonar and Navigation, IEE Proceedings* -, 146(1):2–7, Feb 1999.

[11] L. Chen, S. Shi, K. Bian, and Y. Ji. Optimizing average-maximum ttr trade-off for cognitive radio rendezvous. In *2015 IEEE International Conference on Communications (ICC)*, pages 7707–7712, June 2015.

[12] X. Chen, A. Edelstein, Y. Li, M. Coates, M. Rabbat, and A. Men. Sequential monte carlo for simultaneous passive device-free tracking and sensor localization using received signal strength measurements. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, pages 342–353, April 2011.

[13] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha. FM-based indoor localization. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, MobiSys '12, pages 169–182, New York, NY, USA, 2012. ACM.

[14] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha. Fm-based indoor localization.

In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, MobiSys '12, pages 169–182, New York, NY, USA, 2012. ACM.

[15] A. J. Coulson, A. G. Williamson, and R. G. Vaughan. A statistical basis for lognormal shadowing effects in multipath fading channels. *IEEE Trans. on Communications*, 46(4):494–502, 1998.

[16] G. Deak, K. Curran, and J. Condell. A survey of active and passive indoor localisation systems. *Computer Communications*, 35(16):1939 – 1954, 2012.

[17] A. K. Dey. Understanding and using context. *Personal Ubiquitous Comput.*, 5(1):4–7, Jan. 2001.

[18] Y. Ding, B. Banitalebi, T. Miyaki, and M. Beigl. Rftraffic: Passive traffic awareness based on emitted rf noise from the vehicles. In *ITS Telecommunications (ITST), 2011 11th International Conference on*, pages 393 –398, aug. 2011.

[19] A. Doucet, N. de Freitas, and N. Gordon. Sequential monte carlo methods in practice. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science. Springer New York, 2001.

[20] S. Feldmann, K. Kyamakya, A. Zapater, and Z. Lue. An indoor bluetooth-based positioning system: Concept, implementation and experimental evaluation. pages 109–113, 2003.

[21] B. Ferris, D. Fox, and N. Lawrence. Wifi-slam using gaussian process latent variable models. In *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, IJCAI'07, pages 2480–2485, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.

[22] A. Ferscha, K. Zia, and B. Gollan. Collective attention through public displays. In *2012 IEEE Sixth International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, pages 211 –216, 2012.

[23] D. Figo, P. C. Diniz, D. R. Ferreira, and J. a. M. Cardoso. Preprocessing techniques for context recognition from accelerometer data. *Personal Ubiquitous Comput.*, 14(7):645–662, Oct. 2010.

[24] K. Forster, D. Roggen, and G. Troster. Unsupervised classifier self-calibration through repeated context occurences: Is there robustness against sensor displacement to gain? In *International Symposium on Wearable Computers (ISWC)*, pages 77 –84, 2009.

[25] S. Gupta, K.-Y. Chen, M. S. Reynolds, and S. N. Patel. Lightwave: Using compact fluorescent lights as sensors. In *Proceedings of the 13th international conference on Ubiquitous computing (UbiComp 2011)*, 2011.

[26] S. Haller. The things in the internet of things. In *Proceedings of the Internet of Things Conference 2010*, 2010.

[27] D. Halperin, W. Hu, A. Sheth, and D. Wetherall. Tool release: Gathering 802.11n traces with channel state information. *ACM SIGCOMM CCR*, 41(1):53, Jan. 2011.

[28] M. Hardegger, G. Tröster, and D. Roggen. Improved actionslam for long-term indoor tracking with wearable motion sensors. In *Proceedings of the 2013 International Symposium on Wearable Computers*, ISWC '13, pages 1–8, New York, NY, USA, 2013. ACM.

[29] M. Hazas and A. Hopper. Broadband ultrasonic location systems for improved indoor positioning. *Mobile Computing, IEEE Transactions on*, 5(5):536–547, May 2006.

[30] P. Hillyard, D. Maas, S. N. Premnath, N. Patwari, and S. K. Kasera. Through-wall person localization using transceivers in motion. *CoRR*, abs/1511.06703, 2015.

[31] M. Hossain, M. Hassan, M. Qurishi, and A. Alghamdi. Resource allocation for service composition in cloud-based video surveillance platform. In *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on*, pages 408–412, July 2012.

[32] O. Kaltiokallio, H. Yi?itler, R. Jäntti, and N. Patwari. Non-invasive respiration rate monitoring using a single cots tx-rx pair. In *Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on*, pages 59–69, April 2014.

[33] F. Kawsar, G. Kortuem, and B. Altakrouri. Supporting interaction with the internet of things across objects, time and space. In *Proceedings of the Internet of Things Conference 2010*, 2010.

[34] N. Kern, B. Schiele, and A. Schmidt. Multi-sensor activity context detection for wearable computing. In E. Aarts, R. Collier, E. Loenen, and B. Ruyter, editors, *Ambient Intelligence*, volume 2875 of *Lecture Notes in Computer Science*, pages 220–232. Springer Berlin Heidelberg, 2003.

[35] A. E. Kosba, A. Saeed, and M. Youssef. Rasid: A robust wlan device-free passive motion detection system. *CoRR*, abs/1105.6084, 2011.

[36] A. E. Kosba, A. Saeed, and M. Youssef. Rasid: A robust wlan device-free passive motion detection system. In *Proceedings of the 10th IEEE International Conference on Pervasive Computing and Communications (PerCom2012)*, 2012.

[37] M. Kranz, P. Holleis, and A. Schmidt. Embedded interaction: Interacting with the internet of things. *Internet Computing, IEEE*, 14(2):46 –53, 2010.

[38] J. Krumm. Ubiquitous computing fundamentals. 2009.

[39] K. Kunze, G. Bahle, P. Lukowicz, and K. Partridge. Can magnetic field sensors replace gyroscopes in wearable sensing applications? In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–4, Oct 2010.

[40] J. Lester, T. Choudhury, and G. Borriello. A practical approach to recognizing physical activities. In K. Fishkin, B. Schiele, P. Nixon, and A. Quigley, editors, *Pervasive Computing*, volume 3968 of *Lecture Notes in Computer Science*, pages 1–16. Springer Berlin / Heidelberg, 2006.

[41] A. M.S., M. S., G. N., and C. T. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188, Feb 2002.

[42] K. Muthukrishnan, M. Lijding, N. Meratnia, and P. Havinga. Sensing motion using spectral and spatial analysis of wlan rssi. In *Proceedings of Smart Sensing and Context*, 2007.

[43] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil. Landmarc: indoor location sensing using active rfid. *Wirel. Netw.*, 10(6):701–710, Nov. 2004.

[44] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil. Landmarc: Indoor location sensing using active rfid. *Wirel. Netw.*, 10(6):701–710, Nov. 2004.

[45] G. Ogris, P. Lukowicz, T. Stiefmeier, and G. Gerhard Tröster. Continuous activity recognition in a maintenance scenario: combining motion sensors and ultrasonic hands tracking. *Pattern Analysis & Applications*, 15:87–111, 2012.

[46] K. Ohara, T. Maekawa, Y. Kishino, Y. Shirai, and F. Naya. Transferring positioning model for device-free passive indoor localization. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 885–896, New York, NY, USA, 2015. ACM.

[47] V. Otsason, A. Varshavsky, A. LaMarca, and E. de Lara. Accurate gsm indoor localization. In *UbiComp 2005: Ubiquitous Computing*, volume 3660 of *Lecture Notes in Computer Science*, pages 141–158. Springer Berlin Heidelberg, 2005.

[48] H. Palaiyanur, K. Woyach, R. Tandra, and A. Sahai. Spectrum zoning as robust optimization. In *New Frontiers in Dynamic Spectrum, 2010 IEEE Symposium on*, pages 1 –12, april 2010.

[49] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line. In *Proceedings of the 9th International Conference on Ubiquitous Computing (UbiComp 2007)*, pages 271–288, 2007.

[50] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca. Breathfinding: A wireless network that monitors and locates breathing in a home. *IEEE Journal of Selected Topics in Signal Processing*, 8(1):30–42, Feb 2014.

[51] N. Patwari and J. Wilson. Spatial models for human motion-induced signal strength variance on static links. *IEEE Transactions on Information Forensics and Security*, 6(3):791–802, September 2011.

[52] G. Pirkl, K. Stockinger, K. Kunze, and P. Lukowicz. Adapting magnetic resonant coupling based relative positioning technology for wearable activitiy recogniton. In *Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on*, pages 47–54, Sept 2008.

[53] A. Popleteev, V. Osmani, and O. Mayora. Investigation of indoor localization with ambient FM radio stations. In *Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*, pages 171 –179, march 2012.

[54] Q. Pu, S. Gupta, S. Gollakota, and S. 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.

[55] M. Reschke, S. Schwarzl, J. Starosta, S. Sigg, and M. Beigl. Context awareness through the rf-channel. In *Proceedings of the 2nd workshop on Context-Systems Design, Evaluation and Optimisation*, 2011.

[56] M. Reschke, J. Starosta, S. Schwarzl, and S. Sigg. Situation awareness based on channel measurements. In *Proceedings of the fourth Conference on Context Awareness for Proactive Systems (CAPS)*, 2011.

[57] B. Schilit, N. Adams, and R. Want. Context-aware computing applications. In *Proceedings of the 1994 First Workshop on Mobile Computing Systems and Applications*, WMCSA '94, pages 85–90, Washington, DC, USA, 1994. IEEE Computer Society.

[58] M. Scholz, S. Sigg, H. R. Schmidtke, and M. 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 2011*, 2011.

[59] M. Scholz, S. Sigg, D. Shihskova, G. von Zengen, G. Bagshik, T. Guenther, M. Beigl, and Y. Ji. Sensewaves: Radiowaves for context recognition. In *Video Proceedings of the 9th International Conference on Pervasive Computing (Pervasive 2011)*, 2011.

[60] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-keyi, and M. Youssef. Nuzzer: A large-scale device-free passive localization system for wireless environments. *IEEE Transactions on Mobile Computing*, 12(7):1321–1334, July 2013.

[61] M. Seifeldin and M. Youssef. Nuzzer: A large-scale device-free passive localization system for wireless environments. *CoRR*, abs/0908.0893, 2009.

[62] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka. Precise indoor localization using phy layer information. In *ACM Hotnets*. ACM HotNets, November 2011.

[63] S. Shi, L. Chen, W. Hu, and M. Gruteser. Reading between lines: High-rate, non-intrusive visual codes within regular videos via implicitcode. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 157–168, New York, NY, USA, 2015. ACM.

[64] S. Shi, S. Sigg, L. Chen, and Y. Ji. Accurate location tracking from csi-based passive device-free probabilistic fingerprinting. *Submitted to Transactions on Vehicular Technology, IEEE, Under Major Revision*, 2016.

[65] S. Shi, S. Sigg, and Y. Ji. Activity recognition from radio frequency data: Multi-stage recognition and features. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–6, Sept 2012.

[66] S. Shi, S. Sigg, and Y. Ji. Activity recognition from radio frequency data: Multi-stage recognition and features. In *Proceedings of the 5th IEEE Conference on Context Awareness for Proactive Systems (CAPS2012)*, 2012.

[67] S. Shi, S. Sigg, and Y. 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.

[68] S. Shi, S. Sigg, and Y. Ji. Joint localization and activity recognition from ambient fm broadcast signals. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 521–530, New York, NY, USA, 2013. ACM.

[69] S. Shi, S. Sigg, and Y. Ji. *Wireless Algorithms, Systems, and Applications: 8th International Conference, WASA 2013, Zhangjiajie, China, August 7-10, 2013. Proceedings*, chapter ActiviTune: A Multi-stage System for Activity Recognition of Passive Entities from Ambient FM-Radio Signals, pages 221–232. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[70] S. Shi, S. Sigg, and Y. Ji. Probabilistic fingerprinting based passive device-free localization from channel state information. *Accepted by Vehicular Technology Conference (VTC Spring), 2016 IEEE*, 2016.

[71] S. Shi, S. Sigg, W. Zhao, and Y. Ji. Monitoring attention using ambient fm radio signals. *IEEE Pervasive Computing*, 13(1):30–36, Jan 2014.

[72] S. Sigg, M. Beigl, and B. Banitalebi. *Efficient adaptive communicatoin from multiple resource restricted transmitters*, chapter 5.4. Organic Computing - A Paradigm Shift for Complex Systems, Autonomic Systems Series. Springer, 2011.

[73] S. Sigg, U. Blanke, and G. Troester. The telepathic phone: Frictionless activity recognition from wifi-rssi. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, PerCom '14, 2014.

[74] S. Sigg, D. Gordon, G. v. Zengen, M. Beigl, S. Haseloff, and K. David. Investigation of context prediction accuracy for different context abstraction levels. *IEEE Transactions on Mobile Computing*, 11(6):1047 –1059, june 2012.

[75] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl. RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals. *IEEE Transactions on Mobile Computing (TMC)*, 2013. accepted for publication.

[76] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl. Rf-sensing of activities from non-

cooperative subjects in device-free recognition systems using ambient and local signals. *IEEE Transactions on Mobile Computing*, 13(4):907–920, April 2014.

[77] S. Sigg, S. Shi, F. Buesching, Y. Ji, and L. Wolf. Leveraging rf-channel fluctuation for activity recognition: Active and passive systems, continuous and rssi-based signal features. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, MoMM '13, pages 43:43–43:52, New York, NY, USA, 2013. ACM.

[78] S. Sigg, S. Shi, and Y. Ji. Rf-based device-free recognition of simultaneously conducted activities. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 531–540, New York, NY, USA, 2013. ACM.

[79] S. Sigg, S. Shi, and Y. Ji. Teach your wifi-device: Recognise simultaneous activities and gestures from time-domain rf-features. *Int. J. Ambient Comput. Intell.*, 6(1):20–34, Jan. 2014.

[80] T. Sohn, A. Varshavsky, A. LaMarca, M. Y. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W. G. Griswold, and E. de Lara. Mobility detection using everyday gsm traces. In *Proceedings of the 8th international conference on Ubiquitous Computing*, UbiComp'06, pages 212–224, Berlin, Heidelberg, 2006. Springer-Verlag.

[81] M. Stark, J. Krause, B. Pepik, D. Meger, J. J. Little, B. Schiele, and D. Koller. Fine-grained categorization for 3d scene understanding. In *British Machine Vision Conference (BMVC) 2012*, 2012.

[82] E. Thomaz, V. Bettadapura, G. Reyes, M. Sandesh, G. Schindler, T. Ploetz, G. D. Abowd, and I. Essa. Recognizing water-based activities in the home through infrastructure-mediated sensing. In *Proceedings of the 14th ACM International Conference on Ubiquitous Computing (Ubicomp 2012)*, 2012.

[83] I. T. Union. *Recommendation ITU-R*. 2009.

[84] K. Van Laerhoven and H.-W. Gellersen. Spine versus porcupine: a study in distributed wearable activity recognition. In *Eighth International Symposium on Wearable Computing (ISWC 2004)*, volume 1, pages 142 – 149, 2004.

[85] D. Vasisht, S. Kumar, and D. Katabi. Decimeter-level localization with a single wifi access point. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 165–178, Santa Clara, CA, Mar. 2016. USENIX Association.

[86] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. 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.

[87] R. Want, A. Hopper, V. Falcão, and J. Gibbons. The active badge location system. *ACM Trans. Inf. Syst.*, 10(1):91–102, Jan. 1992.

[88] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3):66–75, January 1991.

[89] L. L. Westeyn, G. D. Abowd, T. E. Starner, J. M. Johnson, P. W. Presti, and K. A. Weaver. Monitoring children's developmental progress using augmented toys and activity recognition. *Personal Ubiquitous Computing*, 16(2):169–191, 2012.

[90] C. Wickens and J. McCarley. *Applied attention theory*. CRC Press, 2008.

[91] J. Wilson and N. Patwari. Through-wall tracking using variance-based radio tomography networks. *CoRR*, abs/0909.5417, 2009.

[92] J. Wilson and N. Patwari. Radio tomographic imaging with wireless networks. *Mobile Computing, IEEE Transactions on*, 9(5):621–632, May 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] F. Wu and B. Hubermann. Novelty and collective attention. In *Proceedings of the National Academics of Sciences*, volume 104, pages 17599–17601, 2007.

[95] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. Ni. Pilot: Passive device-free indoor localization using channel state information. In *Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference on*, pages 236–245, July 2013.

[96] Y. Xie, Z. Li, and M. Li. Precise power delay profiling with commodity wifi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, MobiCom '15, pages 53–64, New York, NY, USA, 2015. ACM.

[97] C. Xu, B. Firner, R. S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, and N. 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.

[98] C. Xu, B. Firner, Y. Zhang, R. Howard, J. Li, and X. Lin. Improving rf-based device-free passive localization in cluttered indoor environments through probabilistic classification methods. In *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*, IPSN '12, pages 209–220, New York, NY, USA, 2012. ACM.

[99] Y. Xu, N. Stojanovic, L. Stojanovic, and T. Schuchert. Efficient human attention detection based on intelligent complex event processing. In *Proceedings of the 6th ACM International Conference on Distributed Event-Based Systems*, DEBS '12, pages 379–380, 2012.

[100] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe. Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking. In *ICMI*, pages 140–145, 2007.

[101] M. Youssef, M. Mah, and A. 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.

[102] M. Youssef, M. Mah, and A. Agrawala. Challenges: Device-free passive localizsation for wireless environments. In *Proceedings of the 13th annual ACM international Conference on Mobile Computing and Networking (MobiCom 2007)*, pages 222–229, 2007.

[103] D. Zhang, Y. Liu, X. Guo, M. Gao, and L. M. Ni. On distinguishing the multiple radio paths in RSS-based ranging. In *Proceedings of the 31st IEEE International Conference on Computer Communications (Infocom)*, 2012.

[104] D. Zhang, J. Ma, Q. Chen, and L. 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.

[105] D. Zhang and L. Ni. Dynamic clustering for tracking multiple transceiver-free objects. In *Proceedings of the 7th IEEE International Conference on Pervasive Computing and Communications (PerCom 2009)*, 2009.

[106] Y. Zhao and N. Patwari. Noise reduction for variance-based device-free localization and tracking. In *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*, pages 179–187, June 2011.

[107] C. Zhe. Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond. Technical report, McMaster University, 2003.

Appendix A

Publication List

A.1 Journal Publications

[1] S. Shi, S. Sigg, W. Zhao, and Y. Ji. Monitoring attention using ambient fm radio signals. *IEEE Pervasive Computing*, 13(1):30–36, Jan 2014

[2] S. Shi, S. Sigg, L. Chen, and Y. Ji. Accurate location tracking from csi-based passive device-free probabilistic fingerprinting. *Submitted to Transactions on Vehicular Technology, IEEE, Under Major Revision*, 2016

A.2 Conference and Workshop Publications

[1] S. Shi, S. Sigg, and Y. Ji. Activity recognition from radio frequency data: Multi-stage recognition and features. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–6, Sept 2012

[2] S. Shi, S. Sigg, and Y. 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

[3] S. Shi, S. Sigg, and Y. Ji. *Wireless Algorithms, Systems, and Applications: 8th International Conference, WASA 2013, Zhangjiajie, China, August 7-10, 2013. Proceedings*, chapter ActiviTune: A Multi-stage System for Activity Recognition of Passive Entities from Ambient FM-Radio Signals, pages 221–232. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013

[4] S. Shi, S. Sigg, and Y. Ji. Joint localization and activity recognition from ambient fm broadcast signals. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 521–530, New York, NY, USA, 2013. ACM

[5] S. Shi, S. Sigg, and Y. Ji. Probabilistic fingerprinting based passive device-free localization from channel state information. *Accepted by Vehicular Technology Conference (VTC Spring), 2016 IEEE*, 2016

[6] S. Shi, L. Chen, W. Hu, and M. Gruteser. Reading between lines: High-rate, non-intrusive visual codes within regular videos via implicitcode. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 157–168, New York, NY, USA, 2015. ACM

A.3 Joint Publications

[1] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl. Rf-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals. *IEEE Transactions on Mobile Computing*, 13(4):907–920, April 2014

[2] S. Sigg, S. Shi, and Y. Ji. Teach your wifi-device: Recognise simultaneous activities and gestures from time-domain rf-features. *Int. J. Ambient Comput. Intell.*, 6(1):20–34, Jan. 2014

[3] S. Sigg, S. Shi, F. Buesching, Y. Ji, and L. Wolf. Leveraging rf-channel fluctuation for activity recognition: Active and passive systems, continuous and rssi-based signal features. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, MoMM '13, pages 43:43–43:52, New York, NY, USA, 2013. ACM

[4] S. Sigg, S. Shi, and Y. Ji. Rf-based device-free recognition of simultaneously conducted activities. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 531–540, New York, NY, USA, 2013. ACM

[5] L. Chen, S. Shi, K. Bian, and Y. Ji. Optimizing average-maximum ttr trade-off for cognitive radio rendezvous. In *2015 IEEE International Conference on Communications (ICC)*, pages 7707–7712, June 2015