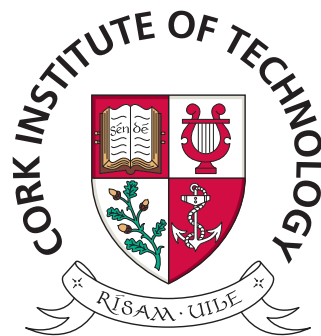# Fall Detection using Channel State Information from WiFi Devices

D. M. Sameera Palipana

Nimbus Centre for Embedded Research

Electronic Engineering, Cork Institute of Technology

Supervised by Prof. Dirk Pesch & Dr. Piyush Agrawal

A thesis submitted for the degree of

Doctor of Philosophy

to Cork Institute of Technology, January 2019

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material to a substantial extent has been accepted for the award of any other degree or diploma of the university of higher learning, except where due acknowledgement has been made in the text.

Signature of Author:

Certified by:

Date:

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dirk Pesch for his mentoring and encouragements during the course of this PhD. I would also like to thank my industry mentor Piyush Agrawal for the guidance and recommendations provided to raise my own standards. I specially appreciate Ramona Marfievici for her support, mentoring and being helpful throughout this research.

I also thank my roommate Indika for bearing with me for four years and his help during the experiments, and lab colleagues David, Victor, Roland, Dylan, Khaled, Pablo, Alejandro and Bastien for the valuable discussions and suggestions.

Most importantly, I would 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.

List of Publications

Conferences:

a) **Palipana, Sameera**, Agrawal, Piyush, and Pesch, Dirk (2016, November). "Channel state information based human presence detection using non-linear techniques." In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys '16), Palo Alto, CA, USA

Journals:

a) **Palipana, Sameera**, Bastien Pietropaoli, and Dirk Pesch. "Recent advances in RF-based passive device-free localisation for indoor applications." Elsevier Ad Hoc Networks 64 (2017): 80-98.

b) **Palipana, Sameera**, David Rojas, Piyush Agrawal, and Dirk Pesch. "FallDeFi: Ubiquitous Fall Detection using Commodity Wi-Fi Devices." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1, no. 4 (2018): 155.

c) **Palipana, Sameera**, Piyush Agrawal, and Dirk Pesch. "FreeFall: A Device-Free Fall Detection System Leveraging Fine-grained WiFi Signals." [Accepted subject to minor revisions], IEEE Journal on Internet of Things (JIOT)

Additional Publications:

a) I. S. A. Dhanapala, R. Marfievici, **S. Palipana**, P. Agrawal and D. Pesch. "Modeling WiFi Traffic for White Space Prediction in Wireless Sensor Networks." IEEE 42nd Conference on Local Computer Networks (LCN), Singapore, 2017

b) I. S. A. Dhanapala, R. Marfievici, **S. Palipana**, P. Agrawal and D. Pesch. "White Space Prediction for Low-power Wireless Networks: A Data-Driven Approach." IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), New York, 2018

Abstract

Falls among the independently living elderly population are a major public health worry, leading to injuries, loss of confidence to live independently and even to death. Each year, one in three people aged 65 and older falls and one in five of them suffers fatal or non fatal injuries. Therefore, detecting a fall early and alerting caregivers can potentially save lives and increase the standard of living. Existing solutions, e.g. push-button, wearables, cameras, radar, pressure and vibration sensors, have limited public adoption either due to the requirement for wearing the device at all times or installing specialized and expensive infrastructure. In this thesis, a device-free, low cost indoor fall detection system using commodity WiFi devices is presented. The system uses physical layer Channel State Information (CSI) to detect falls. Commercial WiFi hardware is cheap and ubiquitous and CSI provides a wealth of information which helps in maintaining good fall detection accuracy even in challenging environments. The goals of the research in this thesis are the design, implementation and experimentation of a device-free fall detection system using CSI extracted from commercial WiFi devices. To achieve these objectives, the following contributions are made herein. A novel time domain human presence detection scheme is developed as a precursor to detecting falls. As the next contribution, a novel fall detection system is designed and developed. Finally, two main enhancements to the fall detection system are proposed to improve the resilience to changes in operating environment. Experiments were performed to validate system performance in diverse environments. It can be argued that through collection of real world CSI traces, understanding the behavior of CSI during human motion, the development of a signal processing tool-set to facilitate

the recognition of falls and validation of the system using real world experiments significantly advances the state of the art by providing a more robust fall detection scheme.

Abbreviations & Acronyms

ADCs Analog-to-Digital Converters

ADL Activity of Daily Living

AoA Angle of Arrival

AP Access Point

BPSK Binary Phase Shift Keying

CFO Carrier Frequency Offset

CFR Channel Frequency Response

CIR Channel Impulse Response

CSI Channel State Information

DACs Digital-to-Analog Converters

DTW Dynamic Time Warping

EMBD Extended Modified B-Distribution

EMD Earth Mover's Distance

FFT Fast Fourier Transform

FFZ First Fresnel Zone

FPR False Positive Rate

GMM Gaussian Mixture Model

HMM Hidden Markov Model

GPS Global Positioning System

ISI Inter Symbol Interference

LoS Line of Sight

LPV Lower Peak Value

LTE Long Term Evolution

MFCC Mel Frequency Cepstral Coefficients

MIMO Multiple Input Multiple Output

MUSIC MUltiple Signal Classification

NIC Network Interface Card

OFDM Orthogonal Frequency Division Multiplexing

PBC Power Burst Curve

PC Principal Component

PCA Principal Component Analysis

QAM Quadrature Amplitude Modulation

QPSK Quadrature Phase Shift Keying

RFID Radio Frequency IDentification

RSSI Received Signal Strength Indicator

SDR Software Defined Radio

STFT Short time Fourier transform

SVM Support Vector Machine

ToF Time of Flight

TPR True Positive Rate

UPV Upper Peak Value

WT Wavelet Transform

WVD Wigner Ville Distribution

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

The share of people over 65 years among the world population is on a steady rise due to increases in human life expectancy. The size of the population of this age group will be close to a quarter of the 20-64 age group by the year 2040 [4]. Each year, about one third of people from this age group are vulnerable to falls and about one out of five people who suffer falls are exposed to head injuries, broken bones in wrists, arms, ankles and hip fractures [5; 6]. Many of those who are not injured, are unable to get up on their own after a fall. This can result in them developing a fear of falling, loss of confidence to live independently, low physical activities, low social contacts, depression resulting in low quality of life and even death [7]. Falls are the leading cause of both fatal and nonfatal injuries for the elderly [8]. Therefore, a system that can automatically detect a fall and inform caregivers or relatives within a short time can be an important health care tool. A reliable fall detection system is helpful to the elderly community in two ways: it reduces the time those who have fallen remain lying on the floor and reduces the fear of falling [7].

A fall detection system typically consists of a detection sensor to detect the fall event and a central unit to alert the caregivers. The sensor and the central unit are connected through a wireless medium and alerting from the base unit is typically done through a landline connection or a cellular service. Existing fall detection sensors can be classified as wearable or ambient sensors [9; 10;

11]. Wearable sensors are either push-button or automatic fall detectors such as accelerometers, gyroscopes, pressure sensors, RFID, and smart phones. A wide range of push button fall detectors, accelerometers, gyroscopes and pressure sensors are currently commercially available[1]. Unlike automatic detectors, push-button systems require the person to be conscious, able and willing to press the button after the fall. Automatic fall detectors measure either human posture or motion related characteristics, such as velocity and movement direction, to infer a fall from a detected event. However, the requirement of always wearing such a sensor may not be followed by elderly people due to forgetfulness or carelessness. For these reasons, research on ambient fall detection systems is currently gaining popularity.

Ambient sensing systems are also called device-free systems because the person being monitored is free from any wearable devices. This sensing technology includes computer vision, microphone arrays, pressure sensors, vibration sensors and RF sensing (WiFi, low power wireless networks and Doppler radar). Computer vision based fall detection approaches are accurate, yet limited in range, fail in darkness and smoke, and can affect occupant privacy [12; 13]. Ambient sensors such as pressure sensors, microphone arrays and vibration sensors can succumb to external sources of pressure, sound or vibrations and can cause false alarms [9; 14; 15]. Especially, vibration sensors have a low range and require a dense deployment for successful deployment [16]. RF sensing is quite a new field compared to the rest. WiFi and Doppler radars are more accurate than low power wireless networks because of the amount of information provided, e.g. high granularity in the signal, Doppler effect and larger bandwidth. Fall detection solutions based on commercial WiFi hardware are much cheaper than Doppler radar, they are ubiquitous and WiFi Channel State Information (CSI) accessed through commercial WiFi chipsets provide comparable features to radar signals while maintaining good accuracy even in large areas. Additionally, such systems can serve a dual purpose by adapting existing WiFi equipment deployed for wireless data communication as a sensing tool to analyze the changes in the radio field to detect human motion. Therefore, this work uses WiFi as the sensing technology for fall detection.

---

[1]Apple Watch 4, Philips Lifeline, Medical Guardian and MobileHelp

1.2 Objectives of the Research

The objective of the research presented in this thesis can be summarized as the design, implementation, experimentation and evaluation of a device-free fall detection architecture using commercial WiFi infrastructure understanding how channel state information is affected by human motion. In particular, the thesis addresses the following key components in the whole system development approach:

- Understand how CSI is affected during human motion and falls through experimentation and mathematical models.

- Develop a CSI signal processing tool chain to facilitate the efficient detection of falls utilizing an understanding of CSI behavior.

- Perform experiments to validate system performance using commercial WiFi infrastructure.

The research therefore approaches the problem of fall detection by understanding, developing and integrating these key components to put in place the whole system solution.

1.3 Approach

The approach is divided into three steps. As the first step, detecting binary occupancy in the observation environment is performed. To detect the occupancy, first, the effect of human motion on the variation of individual CSI subcarrier amplitudes is studied using experiments. Then it is analyzed how the human motion affects the correlations among different subcarriers. Finally, considering the correlations among the subcarriers, a nonlinear approach to detect human presence using CSI provided by off-the-shelf WiFi chipsets is developed. The correlations among subcarriers are exploited by introducing non-linear dimensionality reduction techniques and the human presence is detected with just two WiFi transceivers in the monitoring environment. Similar to existing CSI based human presence detection schemes, this work performs threshold-based detection using

eigendecomposition. Unlike previous work that uses linear techniques for this, the presented solution is derived through exploiting correlations among subcarriers using a nonlinear method. The system was implemented using commercial WiFi devices and experiments were performed in real world environments.

The next step is the design and development of FallDeFi, a fall detection system. FallDeFi provides a novel methodology for time-frequency analysis of CSI for WiFi fall detection including CSI preprocessing techniques. FallDeFi introduces a novel CSI noise filtering and dimensionality reduction approach which helps to construct a spectrogram from CSI amplitude. Novel features are extracted from the spectrogram to detect falls and a feature selection procedure is used to select the features that are resilient to environment changes. A time-frequency based pre-screener is developed for event detection to identify the sections in the spectrogram that are induced by human activities. Performance of the proposed system is evaluated with real world experiments using commodity hardware. The performance is evaluated through experiments using human subjects in four different environments, having different link characteristics (LoS, none LoS, through-wall) and up to two persons using the data collected from commodity WiFi devices. Finally a comparison with two state of the art fall detection systems is performed.

In the final step, FreeFall, the noise filtering and dimensionality reduction steps of FallDeFi are further improved. A combination of in-band noise filtering and a novel principal component selection algorithm removes the artifacts in low frequencies caused by multipath propagation, line of sight attenuation and other interference terms due to signal processing. This helps to obtain precise human velocity profiles from the spectrogram. FreeFall further introduces a time/frequency domain activity segmentation approach to improve fall detection accuracy using human velocity profiles. Finally, extensive experiments are conducted to analyze the impact on fall detection from falling position, distance of the link and room size. The performance of FreeFall is compared with FallDeFi, two other state of the art fall detection systems and an accelerometer-based fall detection system.

1.4   Thesis Organization

The thesis is organized as follows.

- In Chapter 2 a review of the state of the art is provided. The chapter divides human context into three dimensions as space, time and occupants and provides an overview on available state of the art schemes. The granularity of the occupants dimension is further divided as human presence, head count, tracks, activity and identity. The state of the art on fall detection and activity recognition are stated under the same section as they are closely related. The chapter examines both wearable and ambient fall detection methods. Wearable press button fall detectors and wearable accelerometers are currently identified as the most reliable and consequently, they are in commercial production. WiFi fall detection is identified as a new and emerging field which is closely rivaled by established technologies like radar fall detection.

- Chapter 3 provides background theory on topics related to WiFi, Orthogonal Frequency Division Multiplexing (OFDM) and CSI. This chapter further expands on the topic of CSI including what CSI is and how it can be accessed. As CSI cannot be accessed from all commercial chipsets for human activity sensing purposes, this chapter further elaborates the requirements that must be met by commercial chipsets in order to access CSI. This chapter also provides a mathematical model characterizing the behavior of CSI subcarriers when the monitoring area is both free and occupied. This model provides further insights on how CSI is affected by humans.

- Chapter 4 presents the human presence detection system mentioned as the first main contribution of this thesis in Section 7.2. First, the effect of human motion on individual CSI subcarrier amplitudes is analyzed using an experiment. Then the effect of human motion among different CSI subcarrier amplitudes, especially, the linear and nonlinear correlations among the subcarriers is studied. Then, considering these relationships among the subcarriers, a nonlinear approach to detect human presence using CSI provided by off-the-shelf WiFi chipsets is provided. These correlations are

5

exploited by introducing non-linear dimensionality reduction techniques. Human presence is detected with just two WiFi transceivers in the monitoring environment.

- Chapter 5 introduces the next major contribution of this thesis, FallDeFi, the fall detection system comprising of commodity WiFi devices. This system provides a novel methodology for time-frequency analysis of CSI for WiFi fall detection including CSI preprocessing techniques. A pre-screener for fall like events is developed using time/frequency data. Using these techniques, a high performance is obtained compared to the state of the art. Through effective feature extraction and feature selection in time-frequency domains, falls are detected and classified with low dependence on the training environment.

- Chapter 6 presents the final contribution of this thesis FreeFall, a fall detection system enhancing FallDeFi. In FreeFall, the noise filtering and dimensionality reduction steps of FallDeFi are further improved. FreeFall further introduces a time/frequency domain activity segmentation approach to improve fall detection accuracy. Finally, extensive experiments are conducted to analyze the impact of fall position, range from the link and room size on fall detection accuracy.

- Finally, in Chapter 7, the current state of the system is discussed as to how close it is to real world deployment and how it can be improved in future.

Chapter 2

Literature Review

This literature review retains the decomposition of human context proposed by Melfi et al. [1] into three different dimensions of '*time*', '*space*' and '*occupants*'. This decomposition is further illustrated in Figure 2.1. Granularity in *space* can span the whole building, each floor, each room or even a small area in the room. Granularity in time can be either offline, hours, minutes or real-time. Granularity in the *occupants* dimension, which is synonymous to human context recognition, is the most significant information provided by device-free systems. It can be further decomposed as *presence*, *head count*, *tracks*, *activity* and *identity* in increasing order of revealed context or intelligence about the individual. The *time* and *space* dimensions are quite trivial; if a system is capable of providing *occupants* information in *space* in real-time then it is also capable of providing the *occupants* information of rooms, floors or buildings either in real-time or with a reduced time granularity depending on the system's capabilities.

2.1   Human Context Inference

As there is a plethora of different technologies to infer the *occupants* dimension, this chapter only considers the most relevant and significant work to this thesis. Therefore, the state of the art CSI based device-free approaches inferring the *occupants* dimension, i.e., presence detection, head count, tracks, and identity, are presented here. Activity recognition is discussed in depth considering multiple related technologies and fall detection is discussed in conjunction with activity
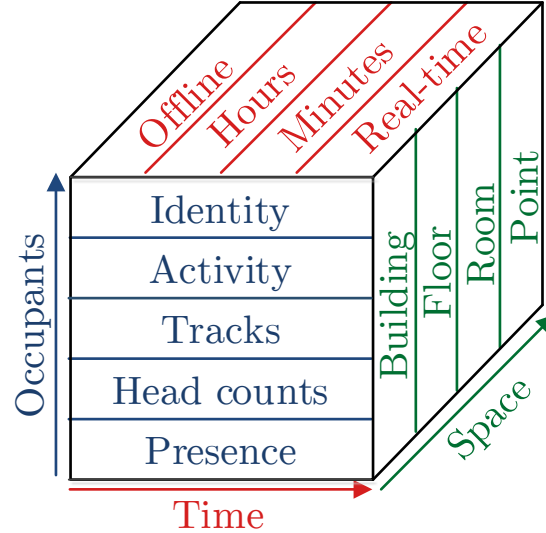
Figure 2.1: Resolution of human context in multiple dimensions proposed in this work. This is a refinement of the occupancy resolutions by Melfi et al. [1]

.

recognition as the signal processing steps in both applications are closely related.

2.1.1 Human Presence

Information whether a space is occupied or not is the most primitive form of intelligence provided by device-free systems. However, occupant presence is a precursor to other forms of human context such as *head count*, *tracks*, *activity* and *identity*. Initially, CSI based solutions were just confined to human motion detection [17; 18; 19]. These systems sense humans through CSI amplitude/phase fluctuations in multiple subcarriers, and these fluctuations caused by humans depend on their motion. In human motion detection using CSI, a distance measure between a CSI fingerprint during human motion and the unoccupied room characterizes human availability. Pilot [20] measured the cross-correlation as the distance measure, and [21] extended it to an adaptive method depending on the environment. In [22] Grey relational analysis was used to measure distance. Omni-PHD [19; 23] used a statistical approach and the fingerprint is a histogram of CSI amplitudes where Earth Mover's distance [24] measures the distance be-

tween unoccupied and occupied histograms. The intuition in these methods is that the affected multi-paths in the wireless channel due to human motion would create enough perturbations in CSI subcarriers to reduce the similarity between unoccupied and occupied fingerprints. The fingerprint obtained is unique to each environment because the multipath characteristics are unique. The fluctuations caused by humans in subcarriers are overlayed in the environment dependent fingerprint and slight changes in the environment require retraining the systems.

Threshold-based approaches reduce the reliance on the environment component by exploiting the correlated fluctuations in subcarriers due to human motion and use a measure of variance to distinguish an unoccupied room from the occupied. Threshold-based approaches exploit the high linear temporal correlation among CSI amplitudes for an unoccupied room. They distinguish the occupied room assuming a loss of correlation in CSI due to temporal amplitude variations. These techniques perform eigendecomposition on the CSI covariance matrix for two reasons, to reduce the number of dimensions and noise in CSI, and to separate the unoccupied room from the occupied with the intuition that eigenvalue variation indicates motion. Note that an eigenvalue of a covariance matrix is a measure of variance and variance in a signal is an indication of human motion.

FIMD [17] first leveraged the correlation among subcarrier amplitudes to detect human *motion*. The authors used first and second eigenvalues to separate the unoccupied room from the occupied using a threshold. They assumed a high linear correlation among subcarriers in the static environment and a low correlation during motion. Using an eigendecomposition of the correlation matrix of CSI, dimensions were reduced and the empty environment was separated from motion. PADS [18] and DeMan [25] extended motion detection to use the phase and amplitude of CSI by assuming linear correlation among subcarriers. They used the first eigenvalue of both amplitude and phase and determined the threshold of unoccupied and occupied scenarios in these two cases by a support vector machine classifier. Although these methods used the first eigenvalue in the detection method, the method later proved to be unreliable due to burst noise [26]. This can reduce the detection accuracy, especially when humans are static. Nonetheless, these schemes focused mainly on detecting a moving person.

Research has recently branched out to stationary person detection from breath-

ing or heartbeat detection [25; 27; 28; 29; 30] or extraction of multiple features from angle of arrival information to distinguish unoccupied environments from occupied [31]. In DeMan [25], a static occupant was detected through their breathing. The method uses the power spectral density of subcarrier amplitudes and isolates frequency components that are related to breathing motion. However, this approach is constrained by the detection range due to transmit power limitations.

Summarizing this section, in general, fingerprinting methods provide good accuracies in cluttered environments, however they are more prone to errors due to changes in the environment that can occur in the real world scenarios and accuracies are sensitive to transmitter receiver positions and the positions of antennas. Thresholding schemes assume that CSI subcarriers are correlated in the empty environment. They also assumed a loss of correlation with human presence. Reduction of CSI dimensions was only performed using linear methods such as eigendecomposition of the CSI correlation matrix. When the human is mostly static, these approaches can have low detection effectiveness. The research introduced in this thesis focuses in a different direction. This thesis explores the effectiveness of nonlinear dimensionality reduction based thresholding approaches for human presence detection.

2.1.1.1 Head Count

Head count is the number of people present in a room or a zone. Therefore, it is essentially a direct extension of human presence detection. Existing solutions capture head counts from fluctuations in the radio signals with the assumption that the fluctuation is proportional to the number of people present [32; 33; 34]. However, this assumption may not hold in real world situations where stationary humans in large spaces can cause high inaccuracies. Therefore, these solutions are mostly confined to smaller spaces, e.g. 4.5 m × 6.5 m [34] and have specific human motion models that emulate a certain motion pattern to cause sustained perturbations in CSI which significantly diminishes the real world applicability.

2.1.2 Tracks

Tracks, commonly known as localization, refers to the trace of where people are at present and where they were before their current position. The important fact is that the subjects are not wearing any wireless device. Therefore, presence and head counts are precursors to tracking. However, tracking can be ambiguous as described by Teixeira et al. [35]. For instance, the track of people leaving the system (e.g. leaving the building) then re-entering the system (e.g. entering the building again, for instance the next day) would not have a track over multiple days if identified only with a temporary and anonymous ID. Therefore, successful tracking require human identities to connect the historical movements with the current.

Tracking using wireless networks can be categorized as RF Tomography, Radio Grids, RF Fingerprinting and RF Backscatter. RF Tomography [36; 37; 38; 39] estimates changes in the propagation field based on the mean attenuation or the shadowing in radio links caused by objects. To compute the attenuation during localization, an offline training phase that is free from humans is a prerequisite. Subsequently, a linearized shadowing model constructs an image of this field [36] or directly provides the object coordinates using particle filters [38]. In essence, RF Tomography shows promise in positioning and tracking of multiple persons (up to 4), in both LoS and non-LoS. It also achieves a low error in the sub-meter range using commercial devices. However, it may be impractical for some applications due to the requirement of high sensor densities in small areas. Furthermore, as the tracked object needs to be surrounded by the transceivers at a height of 1 m above the ground to obtain the object's attenuation image, this sets constraints on transceiver placement.

The distinguishing feature about Radio Grids [40; 41; 42] compared to other approaches is that radio grids utilize models to characterize signal strength fluctuations in multiple links to detect an object's presence and localize it in the target area. Transceivers are merely arranged in grids to extend the target area. These methods largely rely upon a training phase to calibrate the model parameters such as node positions and signal strengths of the links for the unoccupied environment and an online phase to perform the actual localization. Among these works,

LiFS [43] stands out due to the fact that it models subcarrier amplitude variation when the link experiences LoS shadowing by a human as well as fading caused by non-LoS movement. LiFS addresses this by dividing the space surrounding the link into three zones: LoS, non LoS and inside First Fresnel Zone (FFZ) and outside of FFZ and builds three analytical models for the respective zones. This approach is able to locate a human without offline training with 0.5 m median error in LoS and 1.1 m non-LoS. Apart from signal strength information used in the above work, angle of arrival can also be extracted from CSI. Due to different Time of Flight (ToF) of different signal paths, incoming signals encounter varied phases in different antennas as well as among subcarriers of the same antenna. MaTrack [44] exploits this fact to derive the Angle of Arrival (AoA) information using a modified MUltiple Signal Classification (MUSIC) [45] algorithm on each receiver and thereby, uses triangulation to locate a human. The MUSIC algorithm is widely used to estimate the AoA from received signals using multiple antennas. As MaTrack leverages the in-coherency of the reflected path from the human compared to the direct path, it does not require offine training to locate the human and achieves 0.6 m median location error in a 70 m^2 uncluttered area. However, when the person is static, it experiences lower detection rates due to low incoherency of reflected paths.

Tracking by fingerprinting uses the same principle as used in presence detection with an increased complexity which is requiring distinguishability of human locations. A radio map must be initially constructed offline by placing a person at predetermined positions in the area of interest. During the on-line phase, collected CSI values are compared with the fingerprint database, and the corresponding position is inferred. As an example, Pilot [20] used cross correlation as the distance measure to detect occupant. Once an occupant is detected, it uses kernel density based maximum *a-priori* estimation algorithm to estimate the occupant position taking raw CSI values as fingerprints.

Backscattering, the reflection of signals back to where they originated has similarities to the flashing effect in photography or weather radars. It has recently been deployed for human motion tracking as well in which the wireless transmitter acts as the light source. WiDeo [46] extended backscatter measuring to WiFi devices. In WiDeo, wireless transmissions of the WiFi Access Point (AP) is the

equivalent of light source and the reflection of this back at the AP (backscatter sensor) is equivalent to the motion tracing camera. The backscatter sensor distinguishes objects using three features of the reflected RF signal: amplitude, ToF and AoA. The authors built a prototype based on this principle. WiDeo consists of Software Defined Radio (SDR) mimicking the functionality of WiFi APs having a bandwidth of 20 MHz at 2.4 GHz. Each AP is attached to four antennas enabling Multiple Input Multiple Output (MIMO) capabilities which helps with ToF and AoA computations. The authors mention that this system can trace up to five human motions concurrently with a mean error of 12 cm. The impressive fact in this work is the low error achieved in a relatively lower bandwidth scenario compared to other fine grained motion tracking schemes. However, the initial cost for prototype development is $1000-2000. The cost of an SDR primarily depends on its bandwidth, and the bandwidth is related to the overall accuracy.

A concern in common with all technologies that rely on multiple transceivers is the deployment overhead in terms of time and labor effort. This becomes even more complicated in Fingerprinting approaches. In respect to the initial transceiver deployment, RF Tomography and Radio Grids using low power wireless nodes require the most sensing devices, 30-40 for a 70 m^2 area, to obtain good accuracy (accuracy $\leq$ 1 m). RF Fingerprinting and Radio Grids with WiFi devices require fewer sensors, in the range of 5-20 for a similar area depending on the type of radios being used. Hence, deploying such a high density of transceivers in multiple rooms will be demanding in terms of labor and cost. If these devices are battery-powered, maintenance and energy requirements will be additional issues in the long run. Another complication with the deployment of hundreds of transceivers and computation required for processing the data in buildings will be the effect on overall energy consumption of the building. It may turn out for example that the energy savings achievable through the use of occupant detection based climate and energy control may in fact be canceled out by the increased energy usage due to the localization or occupancy detection system.

2.1.3 Activity

Activity refers to what the occupants are engaged in. A type of activity that current systems recognize can range from full body movements such as walking, sitting down, sleeping, standing up and falling as well as activities involving parts of the body such as hand gestures, head movements and mouth gestures etc. Therefore, fall detection is a sub-domain of activity recognition where the system is optimized to distinguish a fall from other Activity of Daily Living (ADL). In Section 2.2 the state of the art of activity recognition and fall detection is discussed in detail.

2.1.4 Identity

Identity verification refers to the process of linking a person with a predefined identity, which can be useful for several applications, such as privacy, security and context-aware services. Systems recognizing human identity can be divided into three categories based on user inputs, what the user carries and user behavioral patterns [47]. The first type refers to the private information of a user such as passwords, pin numbers or patterns. The second type refers to either biometrics of the user or Radio Frequency IDentification (RFID) or magnetic cards carried by them. The third type refers to identifying a person from that person's behavioral patterns. Identity verification using WiFi CSI belongs to the third category. As an example [48] identifies a person from the walking patterns of that person. In this regard this is similar to activity recognition with added complexity where the system has to first recognize the activity and then identify the patterns unique to that person.

2.2 Fall Detection

Existing fall detection sensors can be classified as either wearable or device-free (ambient). Wearable sensors are either push-button or automatic fall detectors. In the following sections these two types of fall detecting sensors are discussed.

2.2.1 Wearable Fall Detectors

Push button fall detection is currently the mainstream commercial fall detection approach. Detectors can be mounted on walls or worn either as pendants around the neck or on wrists [9; 49]. However, unlike automatic fall detectors, if the holder is unconscious, confused or cognitively impaired after the fall, alerting the caregivers becomes a problem.

Automatic fall detectors measure either the human posture or motion related characteristics (velocity, movement direction) to infer a fall from a detected event. Among these approaches, the main technologies are accelerometers [50; 51; 52; 53], inertial sensors [54; 55], barometric pressure sensors and a combination of those technologies [56; 57] and smart phones [58; 59; 60]. Accelerometers, gyroscopes and barometric pressure sensors are at the stage of commercial production and it can be noticed that Apple Watch 4 [61], Philips Lifeline, Medical Guardian and MobileHelp are some of the leading automatic fall detectors [62]. The most common approaches used for fall detection using accelerometers and gyroscopes, especially, are detection of impact with the ground and/or detection of the change in body orientation with respect to a horizontal position. As an example [63] demonstrated that fusing acceleration magnitude, sensor velocity, and body posture features from accelerometers, fall detection accuracy can be further improved. Some algorithms may also detect an inactivity period following the fall, with or without an attempt to recover after the fall [64; 65] this feature helps in reducing the number of false positives which is the main drawback in automatic fall detection systems.

Recently, fall detectors based on smartphones have also gained popularity because smartphones contain multiple sensors, e.g. accelerometers, gyroscopes, magnetometers and Global Positioning System (GPS), and the computational power of smartphones continues to increase. Research shows that a combination of multiple sensors produce the most accurate fall detection. As examples, accelerometer barometer pairs [56] report an accuracy improvement of 11.6% and accelerometer gyroscope pairs [63] report an accuracy improvement of 2.6% compared to the results of using accelerometers alone.

Misclassification of ADLs as falls is a major concern in automatic fall detection

systems. To counter this, some systems offer the option of a feedback system in the form of push button [66] or audio validation [64]. Water resistance of a wearable device is a critical requirement as one third of domestic fall injuries occur in wet conditions [67]. Additionally, the communication range of the fall detection sensor from the central unit and the battery life of the wearable device and central unit are other critical requirements. Some systems also include GPS tracking systems to track the locations in case the subjects move frequently outside the base unit's range. However, the requirement that all wearable sensors must be always attached to the human body may not always be followed by the elderly people due to forgetfulness or carelessness. As an example, one study reports that of people over 90 years that own a push button system, 80% do not press the button, simply because they were not wearing the device when the fall occurred [68; 69]. For this reason, research on ambient fall detection systems is gaining popularity.

2.2.2 Device-free Fall Detectors

Device-free fall detectors, i.e., ambient fall detectors include computer vision, passive infrared sensors, wireless technologies. Computer vision based fall detection approaches have high accuracy [9]. However, they have drawbacks especially in regard to wall penetration, darkness and smoke, and have issues with possible privacy violations. Ambient sensing approaches generally include microphone arrays, air pressure sensors and vibration sensors. These approaches measure changes in the respective sensing environment to detect events and then infer falls from the detected events. However, external sources of pressure, sound or vibrations from the environment can cause false alarms in these systems. Unlike wearable sensors that can be used both indoors and outdoors, ambient sensors work mainly indoors because the infrastructure has to be setup in the building beforehand and they have a limited range. Among the elderly who live independently, about half of the falls occur inside their own premises [70], which indicates the need for reliable indoor fall detection systems.

Wireless technologies include Doppler radars, wireless communication signal descriptors such as CSI and Received Signal Strength Indicator (RSSI) from WiFi

signals, and RSSI from wireless sensor networks. Radars can be accurate, yet they are expensive and have a short range. These two limitations inhibit the deployment of radars in real world applications such as domestic environments. RSSI offered by both WiFi and low power wireless devices are not a good descriptor for fall detection because of low granularity [71]. This is further elaborated in Section 3.2. WiFi CSI based fall detection approaches are quite new compared to the rest of the fall detection sensors, however, they have great potential because the hardware is cheap and ubiquitous and can capture the characteristics similar to radars with good accuracy.

2.2.2.1 Radar-based Activity Recognition and Fall Detection

These approaches detect falls from the shifts in the radar frequencies caused by backscatter due to human motion which is also called the Doppler effect. Radar fall detection has the advantages of privacy preservation, insensitivity to lighting conditions, capability to distinguish motions of animate and inanimate objects like fans and pendulums, and also perform detection, classification, and localization of humans in the target area [12]. Additionally, radars can be used to detect physiological parameters such as heartbeat and respiration of a fallen person [72]. Radars operate in bandwidths in the GHz range using specially built hardware which allows for high time and amplitude resolutions.

A radar fall detection system in its basic form consists of a raw Doppler data collection stage, transformation to a suitable representation such as time domain or time-frequency (e.g. spectrogram), event detection to identify the start and end of an ADL, feature extraction from the ADL and finally classifying the ADL as fall or non-fall. There are three kinds of information provided by radars that can be used to detect falls: Doppler signature, range information and phase and velocity information. Especially, the time frequency signal representation of the Doppler signature is quite similar to time-frequency signal representation of WiFi CSI, therefore, this thesis is inspired by the subsequent signal processing steps used to classify falls from time-frequency signals in radars. Below, alternative solutions for each signal processing step of the Doppler signature in the literature are discussed.

Doppler frequency signatures are analyzed in the time-frequency domain be-

cause human activities captured by radar are non-stationary. Time-frequency representation can capture both frequency content of a fall and location of that frequency content in time. To represent a fall in time-frequency domain, various techniques exist, these can be classified as either linear or quadratic. Short time Fourier transform (STFT) and Wavelet Transform (WT) are the most commonly used linear time-frequency distributions in fall detection. According to the uncertainty principle [73], the time and frequency resolutions of a spectrogram are inversely proportional. Therefore, choosing the optimum STFT window length for the given data has a tradeoff between frequency and time resolutions. WT mitigates this by having high time resolutions in high frequencies and high frequency resolutions in low frequencies. Quadratic representation in detecting falls include Mel Frequency Cepstral Coefficients (MFCC) [74] inspired from sound signal processing, Wigner Ville Distribution (WVD) [75] and Extended Modified B-Distribution (EMBD) [12]. Quadratic time-frequency techniques provide high resolutions, however they can distort the signal due to cross terms. Therefore, they are helpful in visualizing specific signals. As an example, WVD is optimum for linear frequency modulated signals and gives a biased estimate for nonlinear frequency modulated signals [76]. However, recent research comparing features obtained from STFT, WT and MFCC suggest that STFT features perform the best in distinguishing falls [77].

Prior to classifying falls, a fall detection system isolates the beginning and end of an ADL using an event detection scheme. In radar fall detection, the event detection scheme is typically a Power Burst Curve (PBC) [74; 77; 78]. The PBC simply sums the energy between a predefined frequency interval (f_h and f_l) in the spectrogram at time instant t and if the summation exceeds a predefined threshold, this can be considered as an event. The frequency interval f_h and f_l is determined from frequencies that are most common to falls and related activities. As the Doppler frequency shift depends on the operating frequency of the radar, f_h and f_l also depend on this.

The features extracted from radar signal processing include body velocity and acceleration related features of the spectrogram [79]. The commonly captured velocities are torso and limb velocities as torso and limb motions create unique Doppler frequencies that are time varying and together they characterize human

activities. Methods to extract these velocities include percentile method, correlation method and repetition frequency method [79]. Van Dorp and Groen [79] further mention that the correlation method gives the best leg features and the percentile method gives the best torso features. Additionally, the time span of an ADL and the extreme frequency of a fall are other commonly extracted features.

For the classification of falls, binary classification is typically used. The classification tools can be either generative, discriminative or a hybrid of those [13]. As generative classifiers like hidden Markov models can model sequential data, they are popular classification methods in activity recognition. Discriminative models like support vector machines and random forest are efficient at classifying high dimensional data and hybrid models such as Fischer kernels tend to overcome the drawbacks of both types [13].

Apart from Doppler signatures, radars can extract range information of occupants. The idea is that when a person is standing, the signals reflected off that person come from a different distance than while sitting or when a fall occurs. Therefore, the range information changes depending on the person's posture. This idea is exploited by [80] and the authors demonstrate that through the range information extracted from and ultrawideband radar, falls can be distinguished even from fast sitting down which is a typical situation where Doppler radars traditionally fail. However, range information alone can provide false positives, as an example when falling from a lower height, the range information could be similar to sitting down. Therefore, fusion of both Doppler signature and range information from these two types of radars can mitigate the limitations of both.

The phase information contained in the complex high resolution range profile of ultrawideband radars can be used to derive the motion characteristics of a human, including the instantaneous velocity and acceleration. Wang et al. [81] used a Stepped Frequency Continuous Wave radar to extract the range, velocity and acceleration features and argue that these features can be used to detect falls.

The main advantages of radars can be summarized as large bandwidth providing high resolution Doppler signatures and providing features such as range, velocity and acceleration of a falling human. The development of highly customized radar front ends and system on chip solutions require expert knowledge. Therefore, these benefits come at the expense of hardware cost. Coupled with

range limitations, hardware costs restrict the deployment of radars to detect elderly falls in real world situations.

2.2.2.2 RSSI-based Activity Recognition and Fall Detection

It must be noted that not many systems have been developed using RSSI for human activity recognition or fall detection. The main reason being the coarse granularity of RSSI which in turn impacts the accuracy of the activity recognition scheme. RSSI based fall detection approaches can be either model-based [69; 82; 83] or pattern-based [84].

As an example for model-based fall detection, Mager et al. [69] use radio tomographic imaging deploying IEEE 802.15.4 compliant radios. The radios are positioned at two different heights along the boundary of the room. Human pose (stand, sit) is estimated using a hidden Markov model and falls are detected from the time difference in shadowing of the two link levels during the change of the body pose. Kianoush et al. [83] estimates human falls and locations having a grid of IEEE 802.15.4 compliant radios and video cameras on the ceiling. They consider radio links as supplementary to video cameras because changing of the ambient lighting may cause wrong decisions from the cameras. They detect falls by modeling RSSI shifts in radio links due to human activities using a hidden Markov model where hidden states model the temporal sequence of RSSI shifts and state transition probabilities account for time warping effects. Location from radio links is estimated by first building a database of the mean and variance of RSSI for each link for different positions of the target and applying maximum *a-posteriori* estimation for localization. Finally, an extended Kalman filter is applied to fuse location estimates of the cameras and radio links. Harmony [84], a pattern-based approach, extracts time-frequency features from RSSI to classify generic human activities including human falls. They rely on RSSI fluctuations due to signal scattering caused by human motion. Harmony tries to increase the signal to noise ratio by oversampling the RSSI values. They rely on k-means clustering and a Markov model to identify the patterns in the time-frequency domain. Yet accuracy is 5% lower than CSI based activity recognition [26].

It is well known that RSSI fluctuations are coarse-grained due to low resolution (RSSI has 8 bit step size whereas CSI is 64 bit) affecting the overall accuracy [71;

85]. Another issue of using 802.15.4 compliant radios is the low transmission power necessitates the deployment of a high radio density. As an example Mager et al. [69] use a two-level network of 24 RF sensor nodes in a living room in a 3.0 m × 3.0 m arrangement. Hence, deploying such a high density of transceivers in multiple rooms will be demanding in terms of labor and cost. If these devices are battery-powered, maintenance and energy requirements will be additional issues in the long run.

2.2.2.3 CSI-based Activity Recognition and Fall Detection

There exists a plethora of research in CSI based human activity recognition in a wide range of applications. These include multiple activity classification recognition [26; 86; 87], gesture recognition [88; 89; 90; 91; 92], sign language recognition [93; 94], exercise activity recognition [87; 95] or single activity classification such as respiration monitoring [29; 96; 97], sleep monitoring [98] and fall detection [14; 99; 100]. Initial activity recognition schemes such as [26; 86] try to distinguish generic activities such as walking, sitting down, falling, running, brushing teeth, boxing to prove the capability of CSI based activity recognition. Gesture recognition is a subcategory of activity recognition where activity recognition involves both full body and partial movements while gesture recognition considers movements of a certain body part such as hand, head or finger movements. Later, research has branched out to recognize specialized activities such as exercising, dancing, sleep monitoring and fall detection.

Activity recognition. These approaches can be divided into either pattern based, [86; 87] or model based [26; 95; 101; 102]. Among pattern based approaches E-eyes [86] detects generic activities of a single person using location oriented time domain features. E-eyes distinguishes activities as walking or in-place activities using moving variance of CSI. Walking activities are classified using Dynamic Time Warping (DTW) to account for different speeds of walking activities and measuring the distance between training and observation data. In place activities are classified using Earth Mover's Distance (EMD) where CSI histograms of in-place activities are compared. For changes in the link due to environment changes, a semi supervised learning approach is proposed for profile construction based on

user feedback. Especially, this technique in E-eyes inspired this work to extract features resilient to environment changes.

Among the model based approaches, CARM [26] classifies generic activities of a single person including falls using time-frequency features such as torso and limb velocities by characterizing the changing speed of the reflected path length on subcarrier amplitudes. These velocities are quite similar to the velocities obtained from Doppler Radars without the movement direction. CARM models the feature space using a Gaussian mixture model and the state space using a Hidden Markov Model (HMM) and obtains a 96% accuracy for training and testing in the same environment and 80% accuracy for environments and persons it is not trained on. CARM is the closest activity recognition system in the literature to the work in this thesis. The Principal Component Analysis (PCA) based CSI dimension reduction approach used throughout in this thesis is inspired by CARM. CARM introduced the feature extraction from spectrogram to CSI-based activity recognition schemes.

Subsequent activity recognition schemes such as WiFiU [101] and WiDance [95] further improved on these techniques. WiFiU extends CARM to recognize human gait when walking from limb and torso velocities. WiDance extracts Doppler shifts using commodity WiFi devices without any hardware modification to classify exercise game related gestures. They exploit antenna and link diversity to retain relevant Doppler shifts which enables the extraction of motion direction. They achieve 95% overall accuracy which is a 45% improvement over CARM. However, the experimental area is quite limited with the maximum size being 16 m^2 as the direction information in Doppler shifts is corrupted if the person is too far away from the receiver. As an example, the highest accuracy is obtained in WiDance when the human subject is 3 m away. Even though these systems use CSI as the signal descriptor to recognize activities, and have similar signal processing steps in most parts of the system architecture like for fall detection, these systems are tuned to different applications than the application proposed in this thesis. As an example, activity recognition systems which are the closest to fall detection, e.g. CARM, can classify falls among activities that have characteristics that are quite distinct e.g. running, walking, falling, brushing teeth. It is not clear how CARM distinguishes activities with closer characteristics to

falls, e.g. jumping, falling, sitting down and standing up.

Fall Detection. Existing CSI based schemes specialized on fall detection are pattern based and they extract only time domain features [14; 16; 99; 100]. WiFall [14; 99] uses amplitude related time domain features to characterize falls. Extracted features from the amplitude during the detected interval are mainly statistical and these include normalized standard deviation, offset of signal strength, period of motion, median absolute deviation, interquartile range, signal entropy and rate of signal change. WiFall classifies four types of activities, fall, sit, walk and stand and uses multi-class classification to distinguish those activities. From this perspective, WiFall is primarily an activity recognition scheme specialized on fall detection.

RTFall [100] uses both CSI amplitude and phase and builds upon the features extracted in WiFall to increase accuracy. They try to exploit the higher sensitivity of CSI phase compared to amplitude in detecting falls. RT-Fall distinguishes falls from nine additional activities through binary classification. RTFall achieves a true positive rate of 91% and a false positive rate of 8 % which are improvements of 14% and 10% over WiFall respectively. Both RTFall and WiFall are evaluated based on a scenario where a single elderly person is living independently. RTFall reports that the system reacts poorly when there are two people in the target area.

Sensing-Fi [16], tries to mitigate this issue by fusing data from both WiFi and ground vibration detection from accelerometers. As accelerometer ground vibration detection gives false positives when objects fall on the ground, Sensing-Fi mitigates these shortcomings of both technologies and achieves a 95% accuracy after fusion. For fall detection using WiFi, they use the same system as in RT-Fall. However, the experimental area is quite small, which is limited to a 9 m^2 area. They further mention that when the diagonal of the room is more than 3 m, more accelerometers have to be introduced.

One of the main drawbacks of the above mentioned fall detection systems is the low resilience to environment changes. As an example, RT-Fall demonstrates a 33 decrease in true positive rate (the ratio of correctly detected falls out of all the falls) and a 32% increase in false positive rate (the ratio of false alarms out of non falls) when a sofa is moved to the line of sight. Fall detection systems can

be quite specific to the person that the system has been trained to. The above systems do not indicate the impact of accuracy of training on a specific person and adopting to a person with different motion characteristics. Finally, all the above fall detection systems use packet rates of 100 packets/s which translates to a 100 Hz sampling rate. However, activity recognition studies show that sampling rates around 800 Hz must be used to capture all variations that occur in WiFi signals due to human activities [26].

2.3 Summary

CSI-based device free fall detection is cheap, provides high granularity of information, is commercially available and ubiquitous compared to other competitive technologies which provides with an advantage for the deployment in practical environments. Human presence is the most primitive form of human context and this piece of information is important to infer other forms of human context (head count, tracks, activity and identity). Current human presence detection systems using CSI as the signal strength descriptor assume linear correlations among CSI subcarriers and no one has looked at nonlinear techniques to extract features from CSI. Fall detection can be categorized as a specialized form of activity recognition where the system is trained to infer falls from other ADLs. Fall detection using CSI is still at infancy and there are only three main state of the art systems (WiFall, RT-Fall and Sensing-Fi). They all focus on time domain feature extraction. Time-frequency signal processing and feature extraction is popular in Doppler radar and WiFi activity recognition and this thesis is inspired by those systems to develop an environment change resilient and ubiquitous fall detection system.

Chapter 3

Background Theory

WiFi devices that operate according to the IEEE 802.11a/g/n standard use OFDM as the modulation scheme to send data using multiple subcarriers in 20/40 MHz WiFi channels. The receiver measures a discrete Channel Frequency Response (CFR) in time and frequency as phase and amplitude for each antenna pair and report them as CSI. This chapter shows the basic architecture of an OFDM transmitter and receiver used in WiFi and discusses how CSI is extracted from an 802.11n compliant receiver, how the extracted CSI fluctuates during human motion and how these fluctuations can be modeled mathematically.

3.1   OFDM

Wireless signals are subjected to multipath propagation due to reflection, diffraction and scattering off people and objects. Due to multipath propagation, replicas of the original signal with different attenuations and phase shifts interfere with each other at the receiver either constructively or destructively causing Inter Symbol Interference (ISI). In the frequency domain, ISI is translated to frequency selective fading where certain frequencies in the signal experience attenuation and in certain frequencies amplification. Therefore, WiFi uses OFDM which divides a 20/40 MHz channel in multiple subcarriers which effectively transforms a frequency-selective channel into a set of frequency-flat sub channels. As an example, a $20MHz$ WiFi channel is divided into 64 orthogonal subcarriers where each subcarrier can potentially be allocated a different modulation scheme and
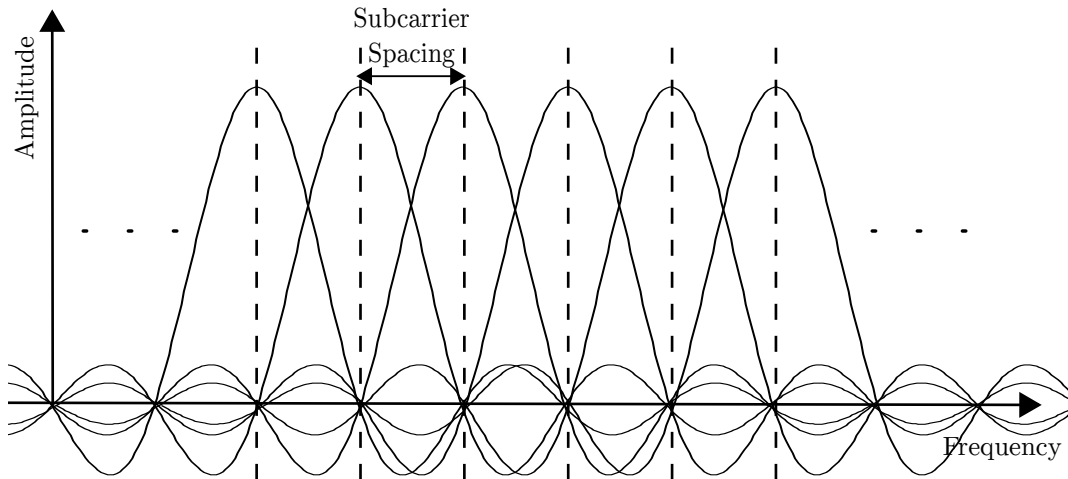
Figure 3.1: A WiFi channel with orthogonal subcarriers.

different data bits. As the subcarriers are orthogonal, the spectrum of each carrier has a null at the center frequency of other carriers and Figure 3.1 illustrates how orthogonal subcarriers are arranged in a WiFi channel. OFDM performs multicarrier modulation for high data rate transmission with efficient utilization of bandwidth and high performance against fading in multipath conditions. Therefore, it is widely used in communication standards such as IEEE 802.11a/g/n, Long Term Evolution (LTE) and LTE Advanced.

Figure 3.2 illustrates the architecture of an OFDM transmitter and a receiver. At the transmitter side, the coded data bits are permutated in such a way that adjacent bits are separated by multiple bits after interleaving. Then the bits are modulated using a digital modulation scheme such as Binary Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK) or Quadrature Amplitude Modulation (QAM). It is possible that different subcarriers may use different modulation schemes depending on the channel conditions. The phase and amplitude of each subcarrier is then fed into a serial-to-parallel converter. Inverse Fast Fourier Transform (IFFT) is then performed to transform from the frequency domain to the time domain and the output samples are then fed to a parallel to serial converter. To reduce inter symbol interference, a cyclic prefix is introduced where the trailing part of the symbol is copied to the front. For the Fast Fourier
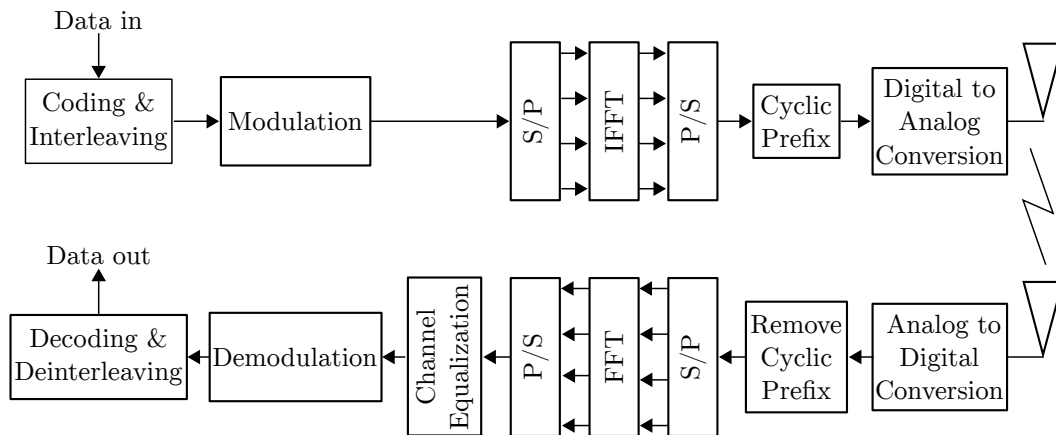
Figure 3.2: 802.11n Compliant transmitter and receiver. [2]

Transform (FFT) algorithm, the cyclic prefix provides an illusion that the signal is periodic. Finally, the real and imaginary components are converted to analogue domain using Digital-to-Analog Converters (DACs) and the analogue signals are then transmitted over the channel.

Upon reception, the receivers sample them and digitize them using Analog-to-Digital Converters (ADCs) and then pass them on to a demodulation process chain. After the serial to parallel conversion, an FFT procedure converts the data samples back into the frequency domain. After parallel to serial conversion, channel equalization is performed to compensate for possible attenuations caused by the channel. Finally, demodulation, deinterleaving and decoding is performed on the data bits.

As CSI extraction is done at the receiver side, the relevant signal processing steps are summarized here. Figure 3.3 illustrates a typical 802.11n receiver with direct down conversion. First the RF signal is amplified by a low noise amplifier. Then it is down converted to in-phase and quadrature components of the baseband signal and sampled by an analog to digital converter. At the packet detection stage, the digitized signal is cross correlated to a 802.11 preamble to detect the incoming packet. Afterwards the Carrier Frequency Offset (CFO) is estimated to compensate for the central frequency offset of the receiver. For each incoming packet, CSI is estimated at the channel estimation block and the subsequent modules compensate for attenuation and phase errors prior to decoding the
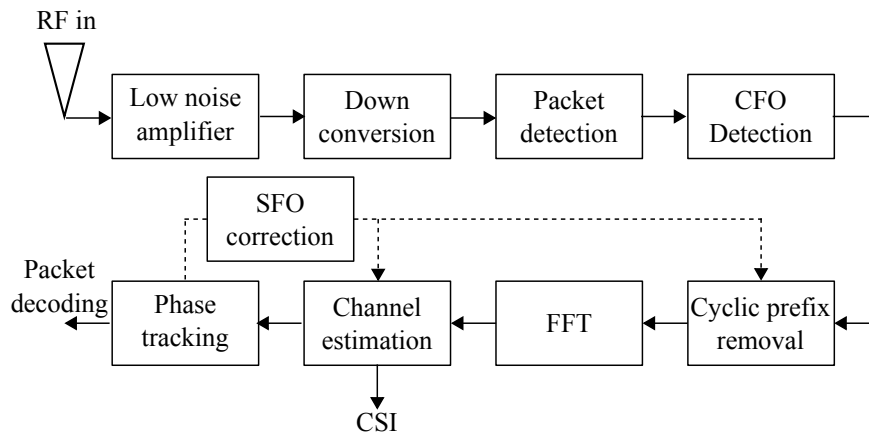
Figure 3.3: 802.11n receiver architecture. [3]

packet. As CSI indicates the influence of the channel on received WiFi packets it can be represented as in Eq. 3.1.

$$Y = HX + n \tag{3.1}$$

where X is the transmitted signal, Y is the received signal, H is the channel and n is noise.

3.2 Channel State Information

CSI represents the channel properties of a communication link in each subcarrier and therefore, provides an estimate of the whole WiFi channel in the frequency domain. The WiFi receiver accurately measures the instantaneous state of the channel using pilot signals that are frequently transmitted over the channel. WiFi receivers estimate the channel at subcarrier level in terms of amplitude and phase and this estimate is fed back to the transmitter as CSI so that the transmitter can adapt its transmissions according to channel conditions.

3.2.1 Comparison of CSI with RSSI

CSI is a better alternative to RSSI because of its higher resolution. As an example, one RSSI value is encoded in 8 bits and one CSI value is encoded in 64

bits. RSSI only provides an average signal strength value for the whole channel, whereas CSI provides the signal strength in the entire 20 or 40 Mhz spectrum as phase and amplitude pairs of 64 subcarriers. This provides additional information about the channel in both time and frequency. When human motion disrupts the WiFi channel, CSI phase and amplitude of each subcarrier also fluctuate which enables the detection of the cause of these fluctuations. Human activities have unique patterns which leave unique footprints in CSI amplitude and phase. This allows pattern recognition schemes to map the human motion with the patterns generated in CSI streams. These patterns can be identified by analyzing a single CSI subcarrier amplitude, phase pair or across multiple subcarriers. As an example of an application that uses multiple subcarriers, the amplitude and phase pairs of multiple subcarriers can be used as multiple sensors to derive the angle of arrival at the receiver with high accuracy [103].

3.2.2 Accessing CSI to infer Human Context

CSI can be measured by 802.11a/g/n WiFi devices that utilize the OFDM modulation for communication. Although CSI is available at the physical layer of the transmitter and receiver internally, it cannot be directly accessed by upper layers for applications such as human motion detection. The main reason for this is that the commercial chip manufacturers do not expose CSI to the users through their drivers and firmware. Currently, this limitation can be overcome in two ways, using commercial chipsets with original drivers and firmware modified by third parties such that CSI is accessible to the users, e.g. Intel 5300 chipset and Atheros 9k series, or using SDRs that consist of the WiFi physical layer.

The advantages of using commercial chipsets over SDRs are that they are much cheaper ($\approx$ \$10 for a commercial chipset and \$2000 for a SDR) and requires less effort to access CSI. Accessing CSI from SDRs require specialized knowledge in integrating the hardware and software and once the motion detection system is developed, the prototype has to be transformed to a system on a chip solution. SDR-based WiFi APs provide very good accuracies in human motion detection because the phase information is not distorted as in commercial chipsets. Phase information is important in activity recognition and localization

schemes to identify Doppler motion direction and angle of arrival of packets. The cost of SDRs are primarily proportional to the bandwidth, and the bandwidth is linked to the overall accuracy of human motion detection. Due to lack of synchronization between transmitter and receiver, commercial WiFi chipsets exhibit issues that curtail the full potential of CSI. As examples, random packet detection delay, sampling frequency offset (phase rotation) and residual carrier frequency offset inherent in these chipsets cause random phase variations over successive received packets. Phase randomness can be eliminated to some extent in the form of a linear transformation (true phase cannot be recovered from this approach) [104; 105]. Nonetheless, recent studies have shown that there are non-negligible non-linear phase errors (due to in-phase and quadrature imbalance issues in direct down conversion) which cannot be eliminated through linear transformation [106]. However, human activity recognition using CSI amplitude alone has shown great success in recent years as researchers are able to understand the behavior of CSI amplitude through mathematical models [26]. Due to these reasons, in this work, only CSI amplitude from commercial WiFi chipsets, i.e. Intel 5300 is used to derive human context from WiFi signals.

3.2.2.1 Characteristics of CSI accessed from Intel 5300 Chipsets

In this work, Intel 5300 chipset was chosen to collect CSI from other commercial chipsets because it is the most commonly used chipset to acquire CSI and therefore has better product support. Although the 802.11n standard specifies that a 20 MHz WiFi channel should be divided into 64 subcarriers and a 40 MHz WiFi channel into 128 subcarries, the Intel 5300 chipset provides CSI in 30 equally spaced subcarriers for both types of channels. Additionally, there are several conditions that have to be met in order to collect CSI: packets have to be transmitted using 802.11n "High Throughput" rate, the receiver's logical configuration and its physical configuration should support the transmission rate. Furthermore, CSI can be accessed by users only after installing a modified driver and firmware from a third party as default driver and firmware do not provide access to CSI.

The Intel 5300 chipset supports four WiFi modes: client mode, access point mode, ad-hoc mode and monitor mode. As monitor mode is the most stable mode out of the four modes in this chipset to extract CSI during long measurement

campaigns, this mode was used for this work. Although monitor mode is used only for observing traffic in a certain channel in general, in this work, specifically, the adapter and driver are configured to "inject" custom frames to be transmitted while the transmitter is still in monitor mode. For CSI collection in monitor mode, both transmitter and receiver should have Intel 5300 chipset installed and the driver and firmware should be customized to access CSI (client mode and access point modes only require one transceiver equipped with this chipset, however, these modes are quite unstable when deployed with the customized driver and firmware).

3.3   Modelling CSI in the Presence of Humans

This section presents a mathematical model for the CSI amplitude.

In wireless radio propagation, radio waves arrive at the receiver over multiple paths due to reflection, diffraction and scattering caused by objects in the environment. Due to this, the received signal is a summation of delayed, attenuated and phase shifted replicas of the transmitted signal that arrive over different paths. The received signal can be characterised using the channel impulse response $h(t; \tau)$ [107],

$$h(t; \tau) = \sum_{i=0}^{M-1} a_i(t) e^{j\phi_i} \delta(t - \tau_i(t)) \tag{3.2}$$

where $a_i(t)$ is the amplitude of the ith path, ϕ_i is the total phase shift caused by physical propagation phenomena such as reflections and diffractions, $\tau_i(t)$ is the time varying delay of ith path and M is the number of paths. Even though this model isolates individual multipaths, due to bandwidth limitations, WiFi receivers can only distinguish clusters of multipath components [108].

The frequency domain representation of $h(t; \tau)$ is the CFR $H(t; f)$:

$$H(t; f) = \sum_{i=0}^{M-1} a_i(t) e^{j\phi_i} e^{-j2\pi f \tau_i(t)} \tag{3.3}$$

WiFi devices that operate in IEEE 802.11a/g/n/ac use OFDM as the modulation

scheme over multiple subcarriers in a WiFi channel to send data. The receiver measures a discrete CFR in time and frequency as phase and amplitude in the form of CSI for each antenna pair:

$$H(t; f_i) = [H(t; f_1), H(t; f_2) \ldots H(t; f_n)] \tag{3.4}$$

where $H(t; f_i) = ||H(t; f_i)||e^{j\angle H(t;f_i)}$ and n is the number of subcarriers.

The purpose of this section is to model and derive expressions for *amplitude* measurements $||H(t; f_i)||$ of CSI subcarriers obtained from the Intel 5300 WiFi chipsets that depend on a person's position. The phase response is not verifiable from the current state of CSI provided by the Intel 5300 chipset. These off-the-shelf WiFi devices produce carrier frequency offsets due to small carrier frequency differences in the transmitting and receiving radios leading to phase randomness in CSI among successive packets [26; 109]. Although there is a calibration method to overcome phase randomness, the actual phase is not recoverable through this approach [110].

The effect of human presence on radio waves can be classified into two groups as *unaffected*: paths that are not affected by the person, and *affected*: paths that have changes in amplitude, phase and delay due to human presence and new paths that are generated due to scattering and reflection off the person. This behaviour is illustrated in Figure (3.4). Hence, the resultant channel in eq. 3.3 can be expressed as just two components, unaffected $H_u(f)$ and affected $H_a(t; f)$ [26; 111]:

$$H(t; f) = H_u(f) + H_a(t; f) \tag{3.5}$$

$H_u(f)$ consists of N unaffected paths. The unaffected paths can be assumed to have fixed lengths, the amplitude a_i and phase $\phi_i - 2\pi f\tau_i$ are time invariant:

$$H_u(f) = \sum_{i=0}^{N-1} a_i e^{j\phi_i} e^{-j2\pi f\tau_i} \tag{3.6}$$

$H_a(t; f)$ consists of $M - N$ affected time varying paths:

$$H_a(t; f) = \sum_{i=N}^{M-1} a_i(t) e^{j\phi_i} e^{-j2\pi f\tau_i(t)}. \tag{3.7}$$
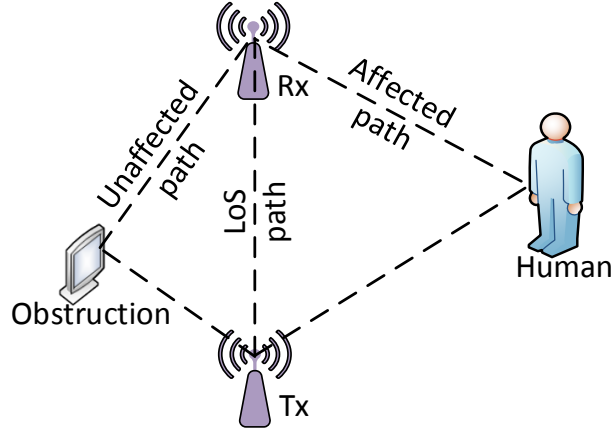
Figure 3.4: The behaviour of affected and unaffected paths due to human influence.

As human movements cause affected path lengths to change, amplitudes and phases of these paths are time varying. This model assumes that there is no other movement in the environment e.g. moving objects.

3.3.1 Modelling the affected and unaffected components

The unaffected component, $H_u(f)$ in this work is modelled as follows. $H_u(f)$ consists largely of the specular LoS paths and the paths resulted from reflections and scattering off static objects. Hence, $H_u(f)$ can be approximated by the CSI observed in the absence of a person in the room, $H_{empty}(f)$ with amplitude $a_{empty}(f)$ and phase $\phi_{empty}(f)$:

$$H_u(f) \approx H_{empty}(f) = a_{empty}(f)e^{j\phi_{empty}(f)} \tag{3.8}$$

Modelling of the affected component, $H_a(f)$, can be done as follows. Due to peculiarities of human bodies, the set of affected paths can consist of a single $N_s = 1$ or multiple $N_s > 1$ specular paths and a collection of diffused paths $N_d \gg 1$ with small amplitudes and random phases. For simplicity, it can be assumed that there is a single specular path with amplitude $a_s(t)$ and delay $\tau_s(t)$, which is reflected by the person. Furthermore, it is assumed that the occupant does not obstruct the LoS paths. LoS obstructions cause high variations in the

subcarriers due to shadowing and diffraction. Therefore, eq. 3.7 reduces as follows:

$$H_a(t; f) \approx a_s(t)e^{-j2\pi f\tau_s(t)+j\phi_s} \tag{3.9}$$

In reality, $H_a(t; f)$ is a time varying result of body scattering and/or reflections that depend on the position of the person. During scattering, it is assumed that the person absorbs the incident power and re-radiates it omnidirectionally. In reflection, the loss is due to only the additional path length and the reflection loss from the object [112]. Therefore, by applying Friis path loss equation, the specular amplitude $a_s(t)$ of the reflected paths can be modelled as [112]:

$$a_s(t) = \frac{kd_0\sqrt{P_{rx}}}{||x_{tx} - x(t)|| + ||x_{rx} - x(t)||} \tag{3.10}$$

where d_0 is the link length, $x(t) = [x_1(t), \ y_1(t)]$ are the coordinates of the person, x_{tx} and x_{rx} are the coordinates of the transmitter and the receiver and k is the reflection loss and P_{rx} is the received power. Due to the person's trajectory, phase shift caused by change of the reflected path length $f\tau_s(t)$ in eq. 3.9 can be written as:

$$f\tau_s(t) = \frac{f \times (||x(t) - x_t|| + ||x_r - x(t)||)}{c} \tag{3.11}$$

where c is the speed of light.

The amplitude response model. From eq. 3.8 and eq. 3.9 the overall channel frequency response $H(t; f)$ can be obtained, and the intended result $||H(t; f_i)||$ by substituting the subcarrier frequencies f_i. Therefore, the amplitude response $||H(t; f_i)||$ of the ith subcarrier with additive white Gaussian noise $\mathcal{N}(0, \sigma_{amp}^2)$ is:

$$||H(t; f_i)|| = \Big[a_{empty}^2(f_i) + a_s^2(t) +$$
$$2a_{empty}(f_i)a_s(t)\cos\left(2\pi f_i\tau_s(t) + \phi_{s,empty}(f_i)\right)\Big]^{\frac{1}{2}} + \mathcal{N}(0, \sigma_{amp}^2) \tag{3.12}$$

The implementation of CSI in the Intel 5300 chipset does not include noise measurements on a per subcarrier basis [113], so a Gaussian distribution with similar parameters, $\mathcal{N}(0, \sigma_{amp}^2)$ can be assumed for all subcarriers. Through CSI amplitudes received during the presence and absence of a person in a real world experiment, the amplitude response of the above model can be validated. The

Table 3.1: Model Parameters.

Parameter	Value	Parameter	Value
TX rate	$1000\,\mathrm{pkts}^{-1}$	TX power	$15\,\mathrm{dBm}$
Human speed	$0.8\,\mathrm{ms}^{-1}$	Duration	$5\,\mathrm{s}$
Link length (d_0)	$5\,\mathrm{m}$	SNR	$100\,\mathrm{dB}$
Reflect. attn.(k)	6	Channel BW	$20\,\mathrm{Mhz}$
Carrier freq.	$5.32\,\mathrm{GHz}$	Antennas	1×1

validation procedure is provided in Section 3.3.2.

3.3.2 Model Validation

To validate the model's behaviour during the presence of a single person, an experiment was carried out where a person moves towards the direct LoS link in a perpendicular trajectory from 4 m away. As shown in Figure 3.5, the person's movement is important in this case to understand its impact on subcarrier amplitudes. Since a constant reflection attenuation (k) is assumed of the reflected path, the best movement to achieve a sufficiently constant reflection attenuation is walking perpendicular to the link where the person's chest is facing the link. The used model parameters are summarized in Table 3.1. The impact of human motion on subcarrier 1 is compared to the real experiment results in Figure 3.6. The top figure corresponds to the model result and the bottom one represents the real fluctuations in the subcarrier.

In the experiment, the person moves towards the link from 4 m away at 0 s. When the person is far from the link, the fluctuations in the subcarrier are small and have a high frequency, whereas closer to the link the fluctuations are large and have a lower frequency. This is caused by the length reduction of the reflected path as the person moves in during this period. The model follows the real pattern until 4.5 s. From there onward, there is a large difference between modeled and real values. This is due to the shadowing of the LoS path by the human which is not modeled in this case.

The amplitude of the real subcarrier has higher variations compared to the modelled amplitude. This is caused by variation of the attenuated reflection

(k) of the reflected path due to movement. In the model, this attenuation is predetermined through a training experiment and assumed constant over time. The other discrepancy between the model and actual results is the phase shift which is a result of a combination of factors such as initial phase of the unoccupied room $\phi_e(f)$, reflections and the fluctuation in human velocity.

In order to quantify the similarity between the real H_r and the modeled H_m CSI matrices, the cross correlation is measured, $-1 < \rho(H_m, H_r) < 1$:

$$\rho(H_m, H_r) = \frac{cov(H_m, H_r)}{\sigma_m \sigma_r} \tag{3.13}$$

where $cov(H_m, H_r)$ is the covariance between H_m, H_r and σ_m, σ_r are the standard deviations of H_m and H_r respectively. Both matrices are of size $(30) \times$ (TX rate$\times 4.5$) where 30 is the number of subcarriers, TX rate $= 1000$ pkts/s and 4.5 is the duration in seconds without the shadowing period. A cross correlation value of 0.9 is obtained between the two matrices. This model is used to understand the parameters that affect the subcarriers due to human influence in Section 5.2.

3.4 Conclusions

This chapter showed the basic 802.11n compliant OFDM transmitter and receiver architecture. The chapter also discussed how CSI can be extracted from
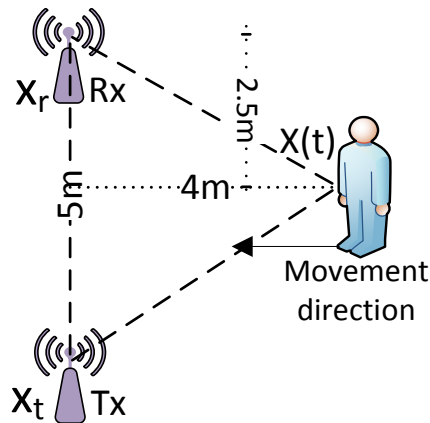
Figure 3.5: Person's position and trajectory during the experiment.
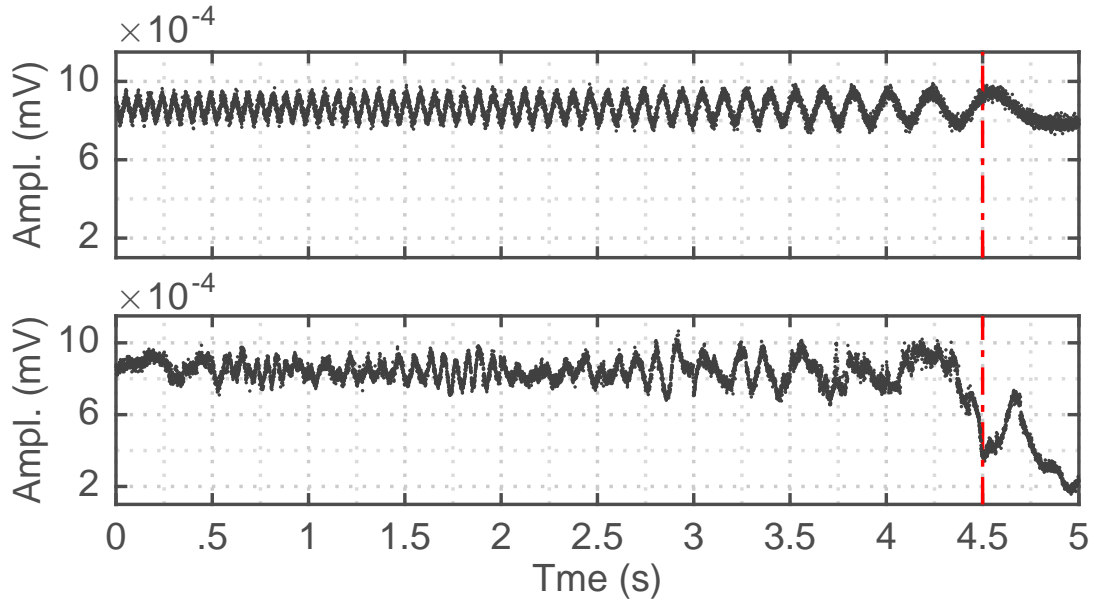
Figure 3.6: Modelled (top) and real (bottom) subcarrier amplitude variation.

an 802.11n compliant receiver and how the extracted CSI subcarrier amplitude fluctuations can be modeled mathematically. Note that the model is developed with the help of existing literature. Especially, dividing the human effects on the channel as $H_u(f)$ and $H_a(f)$ in eq. 3.5 was inspired by [26; 112] and applying the Friis equation for reflected paths from humans as in eq. 3.10 is derived from [112]. The author's own contributions are in equation 3.8 and the verification of the mathematical model using the CSI traces collected from a real world experiment.

Chapter 4

Human Presence Detection

Detecting binary occupancy, or whether there is a person in a particular space, precedes fall detection. Therefore, the system first detects that a person is in the observation environment in order to react to his or her activities. To this end, in this chapter, a time domain event detection scheme is presented.

As mentioned in Section 3.2, for a WiFi data packet received by an antenna pair, a CSI amplitude vector provides 30×1 dimensions. An IEEE 802.11n implementation can use up to 3×3 antenna pairs where dimensions of a CSI data point can amount to 270×1. It is a cumbersome process to train 270 thresholds for each dimension and adjust them when environment conditions change in order to distinguish the unoccupied room from the occupied. Besides, each CSI subcarrier contains noise, which also requires filtering. Hence, a mechanism is required to successfully reduce the dimensions or identify the highest varying dimensions. Additionally, the reduced dimensions also need to contain less noise.

This work exploits correlations among the multiple subcarriers of CSI under human influence and uses kernel PCA to reduce CSI dimensions. Kernel PCA offers adjustable parameters that help in reducing noise to suit different environments. Similar to existing CSI based human presence detection schemes, this work performs threshold-based detection using eigendecomposition. Unlike previous work that uses linear techniques for this, the current solution is derived through exploiting correlations among subcarriers using a nonlinear method. This work is different to the existing research in CSI based human presence detection in three other aspects, *a)* it uses collected data traces to explain the temporal fluctuations

and correlations in subcarriers caused by human motion and then it exploits them to detect human presence, *b)* the dimensionality reduction techniques are parameter-based, so they are adaptable to different environment conditions, and *c)* it focuses not only on detecting human motion, but also a person who is mostly stationary resembling real world situations. In day-to-day life, humans are not completely static or always on the move in domestic environments. Therefore, experiments are conducted to capture such diverse human states.

4.1 Experimental Analysis of CSI under Human Presence

In this section a series of experiments were carried out to explore the temporal behavior of a CSI subcarrier amplitude in a cluttered environment when the room is occupied by a person and when it is not. First the experiments analyze the fluctuations of a single subcarrier. More specifically, the experiments in Section 4.1.1 analyze fluctuations in a single subcarrier during two types of human motion e.g.

- the sensitivity of CSI amplitude to exaggerated human movements such as walking in LoS/non LoS,

- the sensitivity to a stationary person in both LoS and non LoS

in two types of links in different positions of the area of interest. Then the stability of CSI amplitude in the unoccupied environment is explored. If the amplitude is not stable during this state, discerning human motion from unoccupied environment becomes complicated. Then, in Section 4.1.2 the analysis is on fluctuations of multiple subcarriers for the same human motion and later in Section 4.1.3 the analysis is on identifying the correlated fluctuations and their causes. These understandings were necessary to develop the human presence detection scheme presented in this chapter. The experiments were conducted in this way as existing research does not compare the human's motion state, link state and location with the human presence detection accuracy. Therefore, later in Section 4.3 these three parameters (human's motion state, link state and human's location) are used to

evaluate the performance of the human presence detection scheme presented in this chapter.

4.1.1 Temporal Fluctuation of a Single Subcarrier

The experiments were conducted in the 5.32 GHz band so that there is minimum interference noise from other WiFi devices. The transmission rate was 100 pkts/s and the experiments were done in a cluttered environment with tables, chairs and metal drawers. The transmitter and receiver were setup 6 m apart to have a good LoS connection. The transmitter was placed on top of a steel cupboard 1.5 m high. The receiver is on top of a table of 1 m height. In the first experiment, a human walks towards and away from the transmitter within the LoS area. Figure 4.1a shows the experimental setup and Figure 4.1b illustrates the CSI amplitude fluctuation of a single subcarrier under LoS movement of a person. The long rectangles in Figure 4.1a represent tables made of wood and steel and there are two computer monitors scattered on top of each row. It is visible that the amplitude is responsive to such movements and two unique patterns can be observed for both moving in and away from the link.

Figure 4.2b shows the CSI fluctuation of a human crossing the link line in 5 different positions. The fluctuations are higher at *Pos. 1* and *Pos. 5* as shown in Figure 4.2a which are the closest positions to the transmitter and receiver respectively.

Figure 4.3b is an illustration of the movement of a person in the non LoS and Figure 4.3a shows seven different non LoS positions of the person. This experiment shows that the amplitude variation in non LoS movement is not as pronounced as in the LoS movement and fluctuation is also dependent on the position of the room where the human motion occurs. As an example *Area 1* has high variations compared to *Area 6*. This is due to the fact that *Area 1* is closer to the link compared to *Area 6*. Between $0 - 59$ s in Figure 4.3b the temporal variation of the amplitude when the room is unoccupied is shown. The amplitude is stable during this period compared to the rest.

Figures 4.4b and 4.5b illustrate the temporal variations of a subcarrier amplitude for a LoS static person at five different positions and, the response of

a subcarrier for a non LoS static person at 12 different positions. Even though the amplitude variation is higher compared to the empty room for each user position in the two figures, the variations are less pronounced compared to Figures 4.1b,4.2b and 4.3b.

(a)   (b)
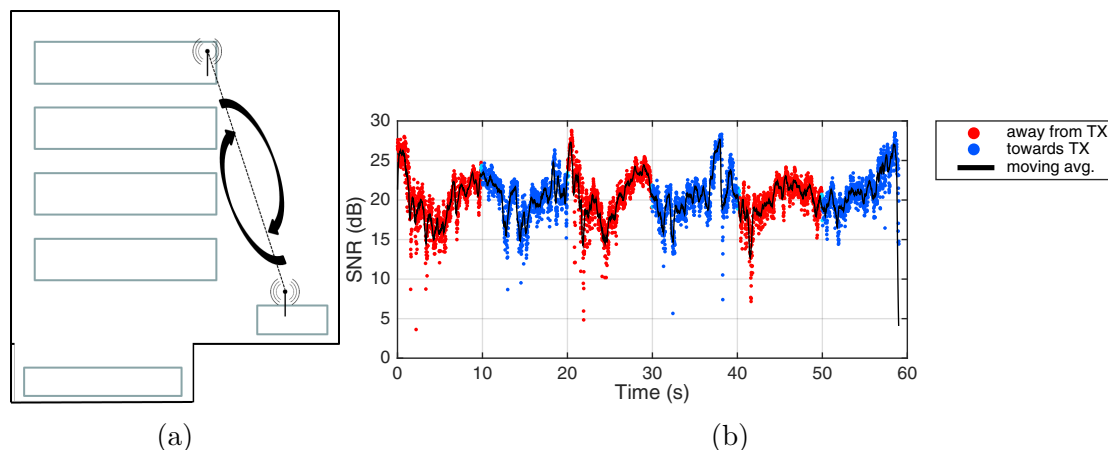
Figure 4.1: Fluctuation of a CSI subcarrier for an occupant moving towards and away from the link. Fig. 4.1a marks the directions of the person in a $5m \times 8m$ room. and Fig. 4.1b shows the variation in a subcarrier for the person's movements.

(a)   (b)
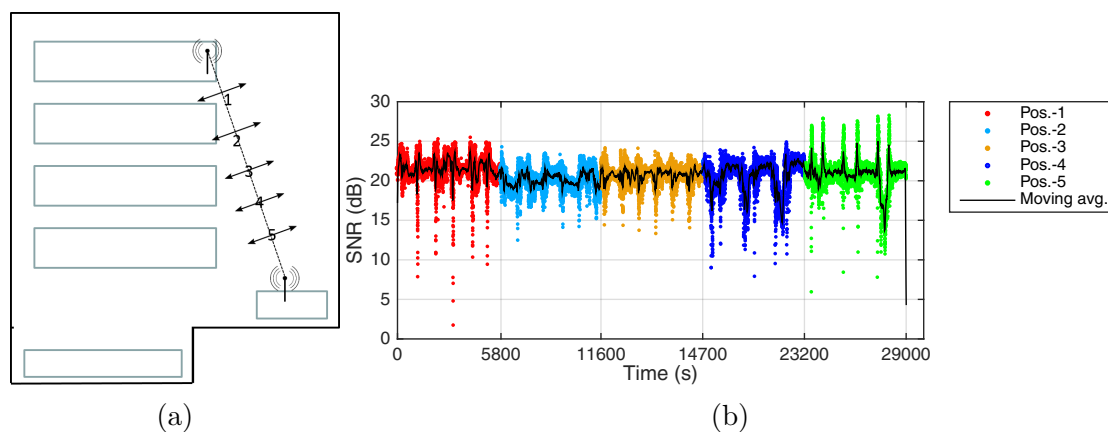
Figure 4.2: Fluctuation of a CSI subcarrier for an occupant crossing the link line at different points. Fig. 4.2a marks the line crossing points of the person in a $5m \times 8m$ room. and Fig. 4.2b shows the variation in a subcarrier for the respective positions.

41

(a)   (b)
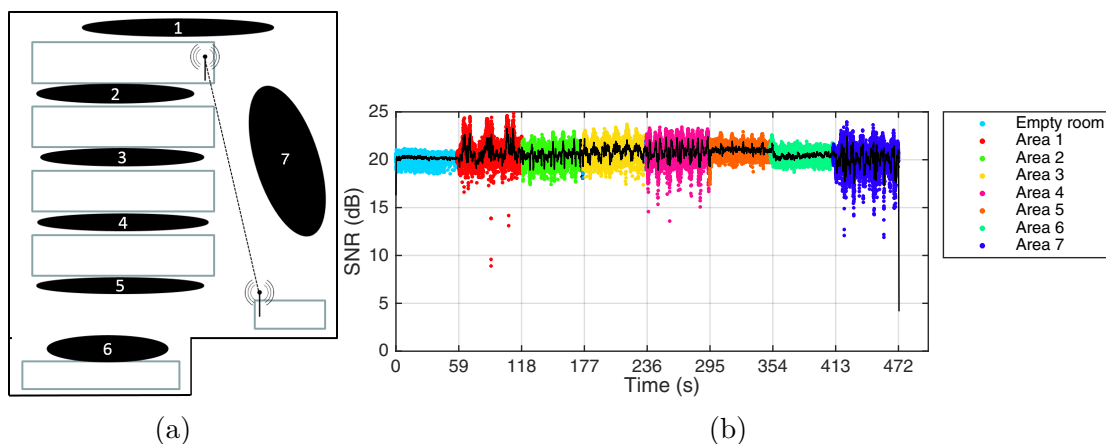
Figure 4.3: Fluctuation of a CSI subcarrier for an occupant moving in the non LoS areas in the room. Fig. 4.3a marks the different positions of the person in a $5m \times 8m$ room. and Fig. 4.3b shows the variation in a subcarrier for the respective positions.

(a)   (b)
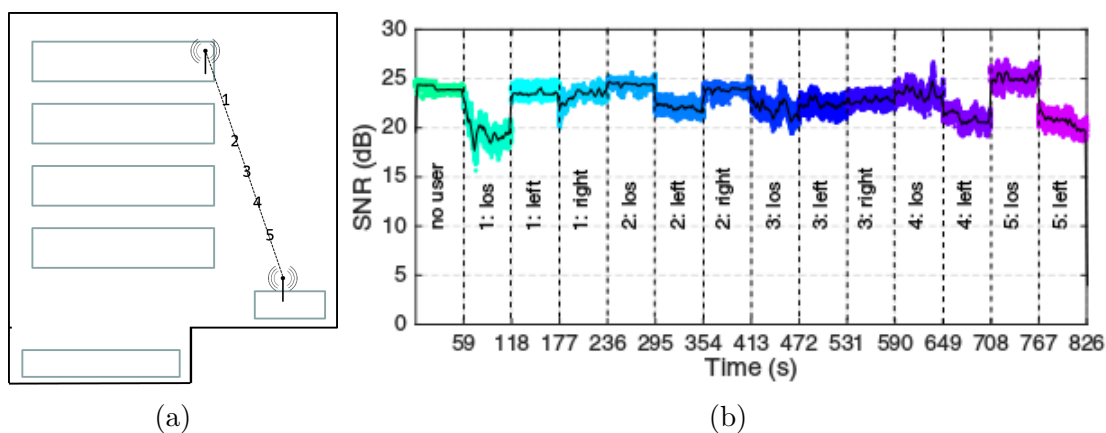
Figure 4.4: Fluctuation of a CSI subcarrier for a LoS static occupant. Fig. 4.4a marks the positions of the person in a 5 m × 8 m room and Fig. 4.4b shows the variation in a subcarrier for the respective positions. 'k: los' indicates the person is in LoS at position k for k = 1,2,..., 5, 'k: left' indicates that the person is 1 m left to k perpendicular to LoS and 'k: right' 1 m right perpendicular to LoS respectively.

42

(a)   (b)
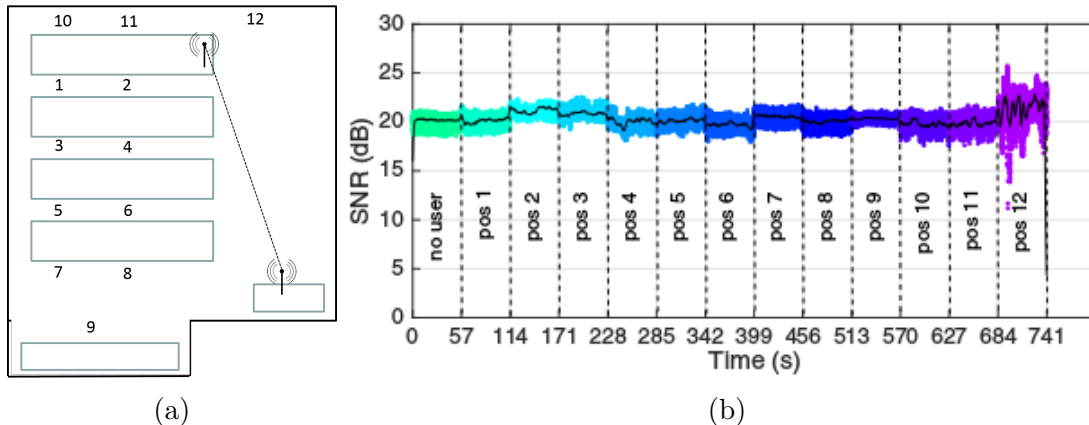
Figure 4.5: Fluctuation of a CSI subcarrier for a non LoS static occupant. Fig. 4.5a marks the positions of the person in a $5m \times 8m$ room. and Fig. 4.5b shows the variation in a subcarrier for the respective positions.

4.1.2 Temporal Fluctuation of Multiple Subcarriers

This section analyzes the behavior of multiple subcarriers when a person is moving in the environment. Figure 4.6 illustrates CSI values acquired from a single antenna pair of the Intel 5300 chipset during a five second interval. An antenna pair provides 30 streams of CSI for 30 subcarriers indexed between -28 to +28 which corresponds to the subcarrier frequencies. The measured subcarrier indices are k = [-28 , -26, -24, -22, -20, -18, -16, -14, -12, -10, -8, -6, -4, -2, -1, 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 28]. As an example, CSI values obtained for the figure correspond to channel 64 having a centre frequency of 5.32 GHz and a bandwidth of 20 MHz.

The person starts to move at 1 s and all the subcarriers begin to fluctuate from 1 s onward. It can be observed that peaks and troughs exist along the axis of the subcarrier index. Even when the person moves from 1 s it can be observed that the peaks and troughs are consistent. This is due to frequency selective fading which is caused by multipath propagation of the WiFi signals. However, the most significant observation from the figure is that adjacent subcarrier fluctuations are correlated to each other, and when the separation between two subcarriers increases, this correlation tends to diminish. Variations in the time axis are predominantly due to human movements and noise existent during the time the

values were collected.
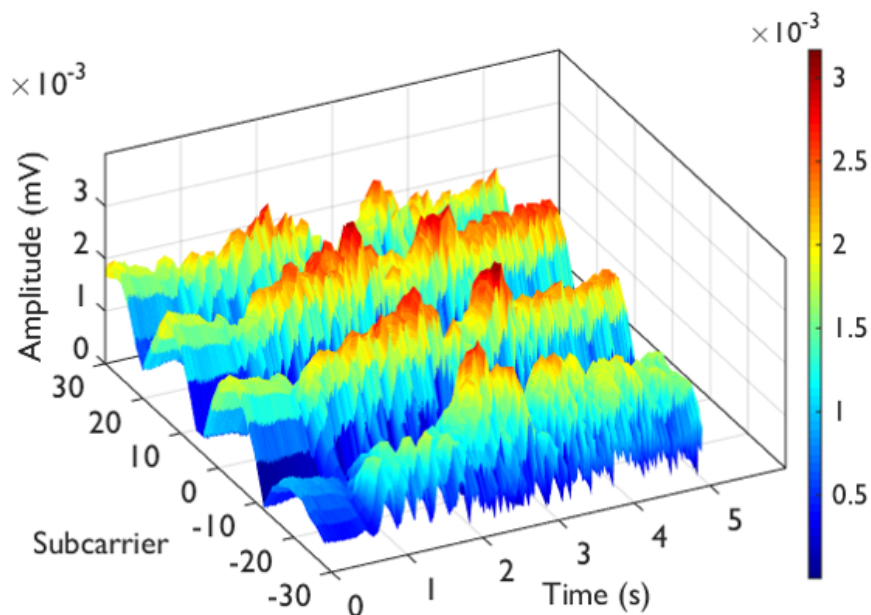
Figure 4.6: A CSI trace from WiFi channel 64 collected during a five second interval.

4.1.3   Correlation of CSI subcarriers

It can be argued that the subcarrier amplitudes have a temporal correlation that originates from several sources. These sources can be identified by analyzing the correlations of CSI subcarriers during a simple experiment. The experiment is conducted as follows. First, the transmitter and the receiver are placed in an empty environment at a five meter distance. Packets are sent at $r_{tx} = 1000$ pkts/s and CSI is collected at the receiver for five seconds. Then a human moves towards the link in a path perpendicular to the link from four meters away and again CSI is collected for five seconds. At first, correlations of subcarriers are studied when the observing environment is empty. Next, subcarrier correlations are analyzed when the person is moving in the environment.

(a)   (b)
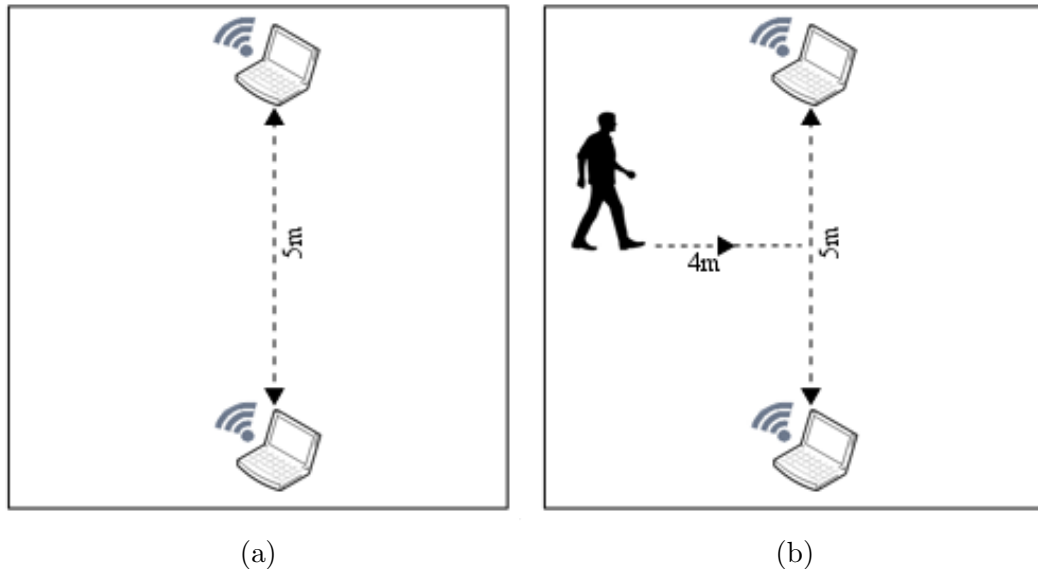
Figure 4.7: Experiment setup of 4.7a) empty and 4.7b) occupied envionments.

These findings can be verified from the data collected during the experiment. As an example, the correlation matrix $\mathbf{H}$ of size 90×90 (90 subcarriers from 1×3 antennas) of the measured data can be calculated as:

$$\mathbf{H} = \mathbf{X}\mathbf{X}^{T} \tag{4.1}$$

where $\mathbf{X}$ is the observation matrix of size $90 \times N_s$. N_s is the number of samples from the experiment that was used to validate the model and it represents the measurement duration t, $N_s = t \times r_{tx}$. The evolution of resultant correlation matrix H with time is illustrated in Figure 4.8 for unoccupied and occupied rooms respectively.

The five figures in the top row in Figure 4.8 correspond to correlation matrices of CSI values collected at $N_s/1000 = 1$ s, 2 s, ..., 5 s durations of the empty environment. The *dark brown* coloured lines that represent correlations less than 0.4 in the five unoccupied room correlation matrices correspond to low correlations among different subcarriers. These low correlations occur when a subcarrier at a peak is compared with a subcarrier at a valley. The valleys and peaks in CSI (as shown in Figure 4.6 along the axis of the subcarrier index) is frequency selective fading due to multipath propagation. Multipath propagation is caused
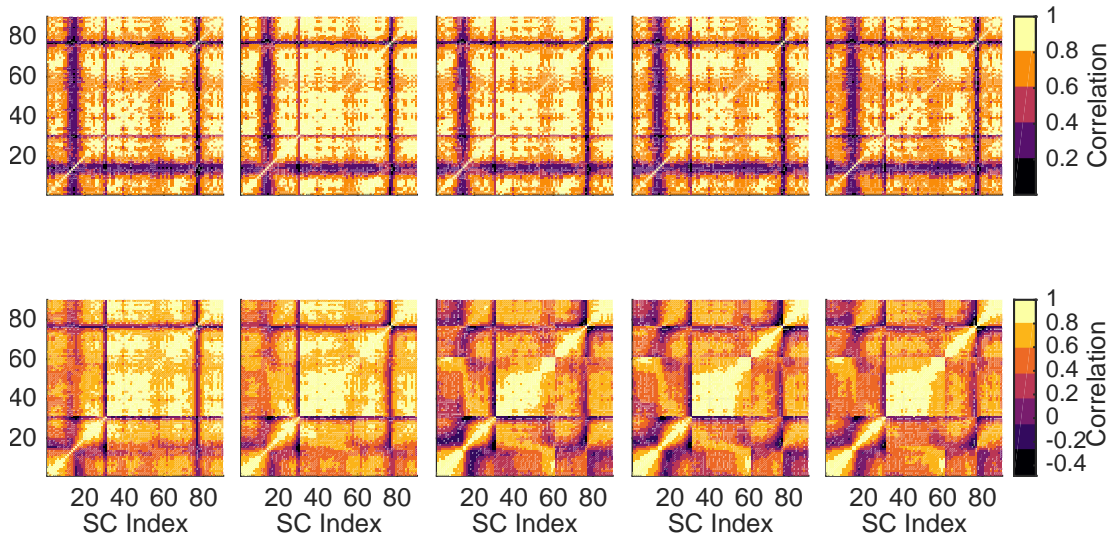
Figure 4.8: The top row represents the correlation matrix of the unoccupied room $\mathbf{H}_{empty}$ and the bottom one corresponds to the correlation matrix with human presence $\mathbf{H}_{occu}$. $N_s = 1000i, 2000i, \ldots 5000i$, where i is the column index.

by scattering of the signal off different objects. Therefore, the more cluttered the environment is, more peaks and valleys can be observed, which result in non-linear correlations among the subcarriers. The top five correlation matrices are similar to each other, this signifies that subcarrier correlation variation is almost constant over the five seconds. This is an expected behavior as there is very little temporal fluctuation in the subcarriers when there is no human motion.

The occupied room is represented by the bottom five correlation matrices in Figure 4.8. The occupied room subcarriers are highly correlated until 2 s ($N_s = 2000$) which is represented by the first two correlation matrices in the bottom row. After 2 s, highest correlation in the occupied room is mostly concentrated among nearby subcarriers of the same antenna (1 - 30 subcarriers: 1st antenna, 31 - 60 subcarriers: 2nd antenna, 61 - 90 subcarriers: 3rd antenna) which corresponds to the light yellow colour along the diagonal in the bottom five figures. Additionally, as N_s increases, the non-linear correlations increase. This is caused by the high temporal fluctuations of the subcarriers as the human moves towards the link. When the human moves closer to the link which is after two seconds, link shadowing causes high amplitude variations in the subcarriers causing high nonlinear correlations.
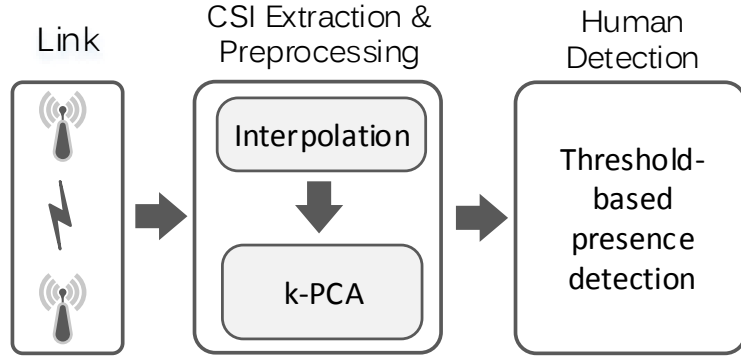
Figure 4.9: Human presence detection scheme. System overview.

The main sources of linear/nonlinear correlations in CSI can be summarized as: *a)* the amplitude and phase of CSI of the empty environment (higher non–linearities can occur in a cluttered environment due to an increase of peaks and valleys in the amplitude response caused by multipath propagation) *b)* human motion, and *c)* measurement duration t. If multiple antennas are used, subcarriers across these antennas also tend to get non-linearly correlated as the measurement duration increases. To a lesser extent, this is true for the subcarriers among same antenna as well. Additionally, amplitude and phase response of each subcarrier is non-linearly correlated.

4.2 Dimensionality Reduction and human presence detection

In this section a system is devised to efficiently exploit the linear/non-linear correlations among the subcarriers that were mentioned in Section 4.1.3 in order to reduce the dimensions, filter noise and detect human presence.

4.2.1 System Overview

The system architecture of the presence detection scheme is illustrated in Figure 4.9. As shown, when CSI is extracted from the packets received through the

---

Algorithm 1: FallDeFi, human presence detection algorithm

    input : CSI_{raw} - CSI measurements of a t_{int} second interval from $m \times n$ antenna pairs

    output: Fall detection

1 **Initialization :** λ_{th} ;

2 **foreach** *3 seconds* **do**

3   **foreach** *subcarrier* **do**

4        $CSI_{interp} \leftarrow$ LinearInterpolate (CSI_{raw});

5   **end**

6   PCs $\leftarrow$ kernel PCA (CSI_{interp});

7   selected PC $\leftarrow$ SelectOptimumPrincipalComponent (PCs);

8   **if** $\lambda_{th} < \lambda_{selectedPC}$ **then**

9   Human present

10   **end**

11 **end**

---

link, CSI streams are first linearly interpolated to account for packet losses and achieve a fixed data rate. Then, nonlinear dimensionality reduction, specifically, kernel PCA is applied to this data and obtain a single temporally fluctuating signal by selecting the optimum principal component and its respective eigenvalue. Finally, presence detection is performed by setting a threshold to this eigenvalue depending on the environment.

4.2.2 Linear Approach: PCA

A method used in the literature to achieve dimensionality reduction assuming linear correlation among subcarriers is PCA [26]. Although PCA is not part of this work, it is explained here briefly because the approach in Section 4.2.3 is a generalization of this technique.

Mathematically, the number of CSI amplitudes N_{obs} obtained from N_{obs} packets can be expressed as a $\mathcal{D} \times N_{obs}$ matrix $\mathbf{X}$, where $\mathcal{D}$ is the number of subcarriers. PCA assumes that rows of this matrix are linearly correlated. It then linearly transforms the rows to a new orthogonal coordinate system. This is done by eigendecomposition of the correlation matrix $\mathbf{K}^{PCA}$ of $\mathbf{X}$:

$$\mathbf{K}^{PCA} = \mathbf{X}^T \mathbf{X} = \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}^{-1} \tag{4.2}$$

Eigenvectors $\{\mathbf{v}_i \in \mathcal{R}^{N_{obs}} \mid i = 1, 2, \ldots \mathcal{D}\}$ form the new directions of the coordinate system. The variances of the new directions are captured by the eigenvalues $\{\lambda_i \in \mathcal{R} \mid i = 1, 2, \ldots \mathcal{D}\}$ where λ_1 has the highest variance. Principal Components (PCs) y_i, where $y_i = \mathbf{K}^{PCA} \times \mathbf{v}_i$ are the coordinates of the new data points. Since the first PC, y_1 contains the majority of variation in the signal, it has a higher signal to noise ratio. Therefore, using the variance/eigenvalue of this PC or another PC that has the signal portion, can be used to separate the unoccupied room from the occupied.

4.2.3 Non-Linear Approach: kernel PCA

Kernel PCA [114] is used in this work to exploit both linear/non-linear correlations in CSI. Kernel PCA first transforms the data to a higher dimensional space and then performs eigendecomposition to reduce the dimensions as in the previous section.

This is achieved by a non-linear transformation of the original data $\{\mathbf{x}_i, \mathbf{x}_j \in \mathcal{R}^{\mathcal{D}} \mid i, j = 1, 2, 3, ..., N_{obs}\}$ into a feature space $\mathcal{F}$ using a function ϕ:

$$\phi : \mathcal{R}^{\mathcal{D}} \to \mathcal{F}, \mathbf{x} \to \mathbf{X}$$

The feature space $\mathcal{F}$ can be of a very high dimensionality. Therefore, without explicitly knowing the transformation function ϕ, Mercer kernels [115] can be used to transform the data [115]. Using a Mercer kernel is analogous to knowing the dot product of the transformed data points $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$, $\mathbf{K} : \mathcal{R}^{\mathcal{D}} \times \mathcal{R}^{\mathcal{D}} \to R$.

Kernels that have been useful in kernel PCA include Gaussian and polynomial kernels [115]. In this work, both these kernels were used to non-linearly transform the data to a higher dimension. Sections 4.2.3.1 and 4.2.3.2 explain the reasons for selecting them and the transformation procedure.

After transforming the data using the selected kernel, eigendecomposition is performed similar to linear PCA to obtain the principal components and the respective eigenvalues. This was implemented according to the procedure mentioned in [116].

After sorting the eigenvalues and respective principal components in descend-

ing order of the eigenvalues, the best *PCs* for human presence detection are chosen. The selection procedure and actual detection is explained in Section 4.2.4.

4.2.3.1   Gaussian Kernel

In this section, the reasons for choosing this kernel and how the parameters were tuned to suit the data are explained. Eq. 4.3 represents the Gaussian kernel used in this work:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{\sigma}} \tag{4.3}$$

where σ is the standard deviation which is a tunable parameter, and $||\mathbf{x}_i - \mathbf{x}_j||$ is the Euclidean distance between two data points $\mathbf{x}_i, \mathbf{x}_j$. This is the most commonly used non-linear kernel in kernel PCA and here the reasons for selecting it are summarized:

  a) Frequency domain representation of the Gaussian kernel is also Gaussian, so it does not introduce high frequency non-linearities. Hence, it provides smooth principal components through transformation [117]. Therefore, when expert knowledge about the non-linearities is lacking, the Gaussian kernel offers a reliable solution under general smoothness assumption.

  b) The width of the Gaussian kernel σ can be adjusted to filter out noise depending on the type of data, this will be explained below.

Parameter optimisation. The standard deviation σ of the Gaussian kernel was optimised for the data as follows.

A set of experiments for two different types of scenarios were conducted, e.g. unoccupied and occupied room as described in Section 4.3, in order to show how σ is determined for the Gaussian kernel. Two data sets of 3000 samples were collected, one each for unoccupied and occupied room experiments. Data transmission rate is $1000pkt/s$, therefore, the 3000 samples correspond to a $3s$ interval. Then the Gaussian transformation of median Euclidean distance of both data sets were plotted in Figure 4.10a against the standard deviation σ. The figure illustrates the variation of $f(\sigma)$ vs. σ where

$$f(\sigma) = e^{-\frac{Md(||\mathbf{x}_i - \mathbf{x}_j||^2)}{\sigma}} \tag{4.4}$$

and $Md(||\mathbf{x}_i - \mathbf{x}_j||^2)$ is the median Euclidean distance of 3000 data points.

The median Euclidean distance is considered instead of the mean because it
is better in describing the central tendencies of the two data sets when the distri-
bution of Euclidean distances are skewed. Unoccupied room data points usually
have a low Euclidean distance because this environment causes low perturbations
in the signal and these perturbations are predominantly caused by noise. If a
fixed σ for the two data sets is used, the Gaussian transformation of the Eu-
clidean distance of the unoccupied room data gives high variations compared to
the occupied. Hence, a mechanism to update σ depending on the nature of the
data is required.

According to Figure 4.10a, it can be observed that median Euclidean distance
$Md(||\mathbf{x}_i - \mathbf{x}_j||^2)$ provides some hint on the nature of the data. According to
eq. 4.3, when σ is high, a data point after Gaussian transformation has larger
influence on neighboring data points leading to low variation and noise in the
principal components. On the other hand, when σ is low, data points have less
influence on neighbour data points and this causes high variation and noise in
the principal components. Hence, depending on the median Euclidean distance of
the data, a threshold, Eud_{th} can be set to separate low Euclidean distances from
high ones so that the low ones correspond to the unoccupied room. Depending on
the threshold, σ can be set to a higher value (saturated region of the unoccupied
$f(\sigma_{unoccu})$ curve in Figure 4.10a) for the unoccupied room to suppress high varia-
tions. Since a high σ value for a data point has larger influence on neighbour data
points, it cancels out the noise. To highlight variations in the occupied room, σ is
set to a value (in the linear region of the occupied $f(\sigma_{occu})$ curve in Figure 4.10a)
for Euclidean distances exceeding Eud_{th}.

4.2.3.2 Polynomial kernel

Eq. 4.5 describes the polynomial kernel used in kernel PCA where c and d are
two adjustable parameters:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + c)^d \tag{4.5}$$

The linear kernel used in PCA is a subset of the polynomial kernel. The poly-
nomial kernel can be transformed to either linear or a higher order polynomial
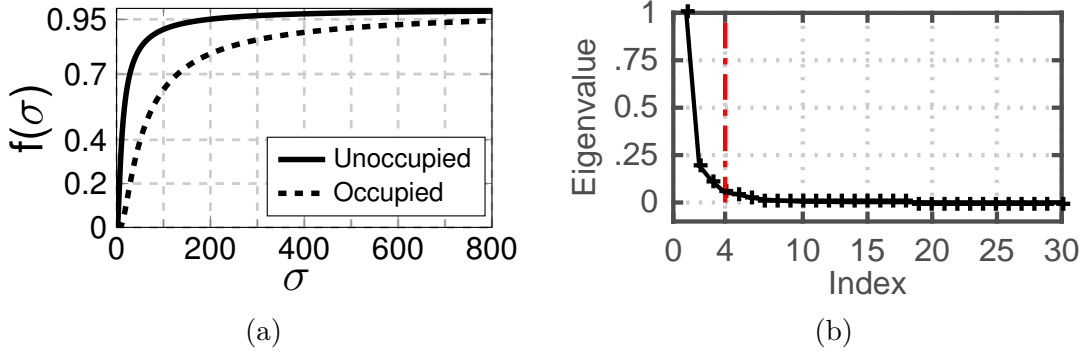
Figure 4.10: Fig. 4.10a illustrates $f(\sigma)$ vs σ of 3000 data points comparing unoccupied and occupied rooms and Fig. 4.10b illustrates a scree plot.

depending on the type of data by adjusting the parameter d in eq. 4.5. Higher order polynomials amplify higher amplitudes in the signal and suppress lower amplitudes such as noise. However, unlike the Gaussian kernel, a polynomial kernel can distort the principal components through the transformation. Adjusting the parameters such as c and d, these distortions can be minimized. It is shown below how the parameters are selected for the polynomial kernel in the current approach.

Parameter selection. For the polynomial kernel in eq. 4.5 there are two parameters that can be adjusted (c, d). For the choice of d, $d = 2$ was kept so that this becomes a second order polynomial. For higher order polynomials very high eigenvalues and principal components can occur especially when the occupant is moving. The computation time of higher order polynomials is also high. Therefore, eq. 4.5 reduces to:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + c)^2 = (\mathbf{x}_i^T \mathbf{x}_j)^2 + 2c\mathbf{x}_i^T \mathbf{x}_j + c^2 \tag{4.6}$$

When the parameter c is high there is increased weight for lower order values, e.g., $2c\mathbf{x}_i^T \mathbf{x}_j$ which provides a smooth signal for both occupied and unoccupied rooms. When c is low, the weight on higher order values is high resulting in high peaks for the principal component when the occupant is moving. In this case, $c = Md(\mathbf{x}_i^T \mathbf{x}_j)$ is set for data points of size N_{obs} so that a balanced signal without unnecessary peaks is achieved.

(a) PC 1   (b) PC 2
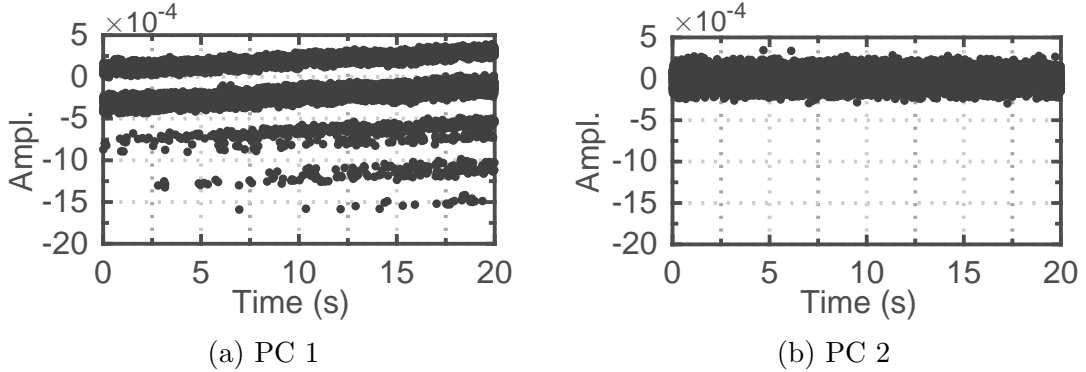
Figure 4.11: PCs of the empty room.

4.2.4   Human Presence Detection

Section 4.2.3 identified the highest varying principal components from the eigen-decomposition of Gaussian and polynomial kernel transformed data. The purpose of this section is to determine the best PC and the respective eigenvalue and determine a threshold to detect human presence.

After applying kernel PCA, as mentioned in section 4.2.3 more dimensions than the original ones can be obtained. It is not known how many Principal Components (PCs) contain parts of the original signal and how many Principal Component (PC)s are just noise. To isolate the signal part from the noise, a scree plot is used as in Figure 4.10b. It plots the amplitude of each eigenvalue with its index. According to the elbow method [118], the point where the steep curve ends (marked with a red dashed line in the figure) is considered to be the last point where the signal portion is carried. For the particular case in the figure, the first four *PCs* can be selected as the best. An eigenvalue of a principal component represents the variance in that PC. When there is human motion, it can be considered that there is high variance in the signal and the resultant principal components. The scree plot helps to select the principal components with the highest variance using the elbow criterion. If one plots the percentage of variance explained by the PCs against the number of PCs (eigenvalue vs PC index in Figure 4.10b), the first four PCs will add much information (explain majority of the variance), but after 4th PC in the figure the marginal gain drops, giving an angle in the graph.

In both PCA and kernel PCA, $PC1$ corresponds to the highest eigenvalue and captures the highest variance of the signal. However in CSI, the first principal component is unusable due to high noise. The reason is that $PC1$ of the unoccupied room has high variation which reduces the overall signal to noise ratio of the occupied and unoccupied rooms. As an example, Figure 4.11 shows the temporal fluctuations of PCs 1 and 2. $PC1$ indicates a signal with multiple levels whereas $PC2$ consists of a single level. It can be argued that these multiple levels are a result of multiple peaks and valleys of CSI caused by frequency selective fading. Although the mean of each subcarrier is subtracted prior to applying kernel PCA, the mean subtracted CSI contains high variation in the subcarriers with peaks and low variation at valleys leading to multiple levels in $PC1$.

Therefore, for human presence detection, the best PC out of the above four PCs are chosen based on the maximum eigenvalue ratio Λ_{max} of empty and occupied scenarios:

$$\Lambda_{max} = \max_k \frac{\lambda_{k,occu}}{\lambda_{k,empty}} \ where \ k = 1, 2, \ldots m \qquad (4.7)$$

PCs 2 and 3 tend to maximize this ratio due to high signal power and low noise in those PCs. In Section 4.3.3.1 it is shown that the performance of PCs 2 and 3 vary depending on the noise profile of the environment.

After selecting the best principal component for detection, separation of the empty scenario from an occupied scenario must be performed for the test data. This is done by setting a threshold for the training data obtained during the absence of a person in the unoccupied room. This threshold λ_{th} corresponds to the highest eigenvalue of the *selected* principal component from n intervals of the training period:

$$\lambda_{th} = \max_i (\lambda_{i,empty}), \ where \ i = 1, 2, \ldots, n \qquad (4.8)$$

If an eigenvalue of test data exceeds this threshold, it is counted as detection of an occupant.
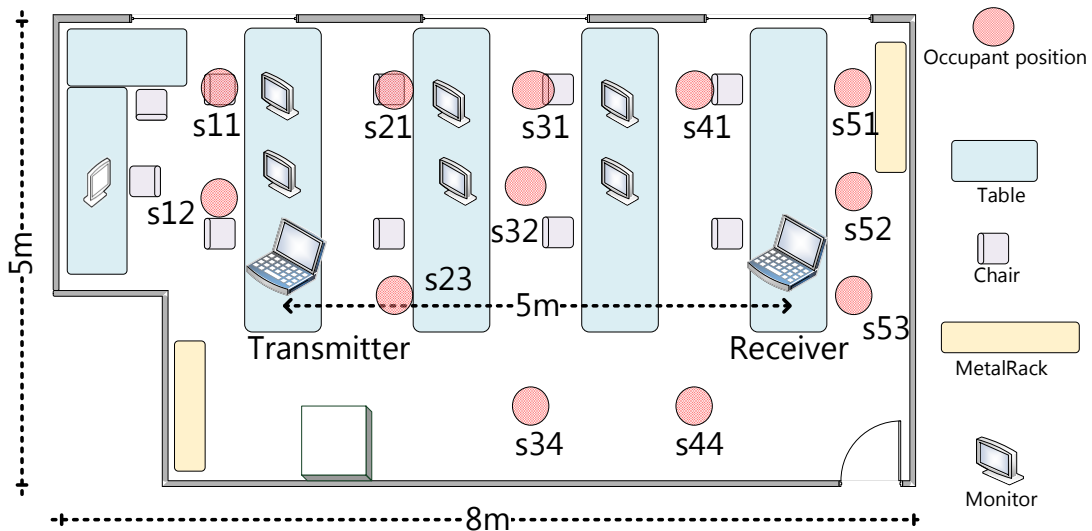
Figure 4.12: Floor plan of the lab. The two laptops operate as the transmitter and the receiver.

4.3 Evaluation

In this section the experimental environment, execution of the experiments and the obtained results are explained.

4.3.1 Experimental Setup

The experiments were performed at a research lab in Nimbus centre and the floor plan is given in Figure 4.12. Data was collected from two experiments LOS, N-LOS corresponding to the link types between the transmitter and receiver, *i.e.,* a) LoS, b) non LoS. In experiment N-LOS, the LoS is blocked by multiple 15 inch desktop computer screens contributing to a ≈ 2.26 dB mean attenuation compared to the LoS link. In each experiment, data was collected for three scenarios: a) unoccupied room (serves as reference data), b) a stationary person, and c) a walking person.

For the experiments (both LoS and non LoS) involving a stationary person, 12 different positions were chosen as indicated in Figure 4.12. In the figure, positions are marked with s_{ij} where i is the row number and j is the column

number. During this experiment, the person stays in each position and is involved in subtle movements with regular pauses. The idea behind this type of movement is to emulate a person a working in a real world scenario as close as possible. 12 different positions were selected as the location of the person in relation to the link has an effect on the detection accuracy due to coverage of the link. Therefore, to measure the effectiveness of the presence detection scheme throughout the area of interest 12 positions covering the entire area were selected. In the experiments with a walking person, two scenarios are considered. In one of them, the person is not obstructing the LoS link (this scenario is called *ldn*: LoS dynamic) and in the other, the person's path crosses the link (*ndn: non LoS dynamic*).

The transmitter and receiver of the experimental setup are two laptops, ACER 5740 and Lenovo Thinkpad L520. Both run Ubuntu version 14.04. Each of them is equipped with three external antennas. A WiFi Network Interface Card (NIC) with the IEEE 802.11n IWL 5300 chipset, modified drivers and firmware are installed in the two laptops in the two laptops to function as the transmitter and receiver. Both WiFi cards in the two laptops were programmed to operate in monitor mode, which is one of the WiFi modes provided by the IEEE 802.11n standard to sniff packets in a particular channel. CSI Tool [119] is leveraged to analyze the data received from the chipset.

4.3.2 Execution

In this section it is explained how the data was collected from the two experiments and how the dimensionality reduction was performed and human presence detection using those data.

Altogether there were 32 measurement sets gathered for experiment LOS and N-LOS, each experiment consisting of 16 scenarios (12 static, 2 dynamic and 2 reference). The two reference data sets were used to calculate false positives. For each scenario, the transmitter was programmed to send packets for 60 s at 1000 pkts/s using injection scripts provided by the CSI Tool. Hence, data is collected for a total of 16×2 scenarios amounting to a total duration of 32 minutes. To avoid interference from other devices, all measurements were collected in channel 64 which is a 20 MHz WiFi channel with centre frequency 5.32 GHz.

The parameters used in the experiments are provided in Table 4.1.

The transmission rate of 1000 pkts/s leads to a two dimensional observation CSI matrix of size $(30 \times N_{tx} \times N_{rx}) \times (N_{obs})$ where $N_{obs} = 60000$ for each scenario. If the observation matrix is directly transformed to the kernel matrix which is of size $N_{obs} \times N_{obs}$, applying PCA or kernel PCA requires heavy usage of system memory as N_{obs} is too large. If N_{obs} is too small, the correlations in the data is lost. Therefore, each scenario of 60 s is divided into 20 continuous subintervals of 3 s containing a batch of $N_{obs} = 3000$ packets. Then kernel PCA was performed separately on the 20 subintervals and obtained 20 pairs of principal components and eigenvalues. There is also a practical reason for choosing $N_{obs} = 3000$. That is, 3000 packets correspond to 3 s which is a good response time for occupancy detection.

Thereafter, for each batch of 3000 packets, kernel PCA is performed with Gaussian and Polynomial kernels to determine the eigenvalues, best principal components, and calculate detection percentages. For the two environments LOS and N-LOS, the parameters for the Gaussian kernel are set as explained in Section 4.2.3.1 as $Eud_{th,A} = 32$, $Eud_{th,B} = 35$ respectively, $\sigma_{unoccu} = f^{-1}(.95)$ and $\sigma_{occu} = f^{-1}(.7)$. A detection percentage for an experiment is $\frac{N_{det}}{N_{int}} \times 100\%$ where N_{det} is the number of intervals that exceeds the detection threshold and N_{int} is the total number of intervals. The detection threshold for a particular environment LOS or N-LOS is determined from the two reference datasets (unoccupied room CSI traces) for the two environments. As each reference dataset consists of 60000 packets corresponding to 60 s, the two datasets are divided into $i = 20$ groups of 3000 packets and the group that has the highest eigenvalue $max(eig_i)$, where eig_i is the eigenvalue of group i is selected as the threshold. Finally, as a comparison, PCA is applied to the same data and the detection percentages for the selected principal components are compared in the following section.

4.3.3 Results Analysis

In this section, the True Positives (TP) and False Positives (FP) of the presence detection scheme are quantified. Furthermore, the performance of the underlying dimension reduction techniques for static and dynamic persons is assessed. Then

Table 4.1: Experiment parameters for the selected scenarios.

Parameter	Value	Parameter	Value
Environment	Cluttered	Antennas	3×3
Exp. duration	60 s	TX power	15 dBm
# of exp.	15	Channel BW	20 MHz
Link length	5 m	Pkt TX rate	1000 pkts/s
Carrier freq.	5.32 GHz	N_{obs}	3000

Table 4.2: Average detection percentages for Exp. LOS and N-LOS.

	PCA		G-kPCA		p-kPCA	
	TP	FP	TP	FP	TP	FP
LOS ($PC\ 2$)	88.3%	0%	90.6%	0%	62.6%	0%
N-LOS ($PC\ 3$)	83.3%	5%	88.3%	0%	62.6%	5%

the impact of the selection of principal component and observation interval (N_{obs}) on overall detection accuracy is evaluated.

4.3.3.1 Overall detection performance

For both Gaussian kernel PCA and PCA in experiment LOS the best performing principal component in terms of detection is $PC\ 2$. For polynomial kernel PCA, $PC\ 3$ has the best performance. However, in experiment N-LOS, $PC\ 3$ is consistently providing good detection percentages. It can be argued that the reason for this is the noise profile of the empty scenario in the two experiments. Since the link in experiment N-LOS is non LoS, the received signal is more noisy than in experiment LOS, which is a LoS link. Hence in experiment N-LOS, $PCs\ 1$ and 2 capture this variance caused by high fluctuations in the link. Therefore, $PC\ 2$ is used for human detection in experiment LOS and $PC\ 3$ for detection in experiment N-LOS. In Table 4.2, the average percentages of true and false positives of all the 14 scenarios are provided for the two experiments. On average, the Gaussian kernel has a better detection performance because of high TPs and low FPs. The reason for false positives to be zero in most cases is that the threshold

λ_{th} is set as the highest eigenvalue from the intervals of the reference unoccupied room data set as explained in eq.4.8.

(a)

(b)
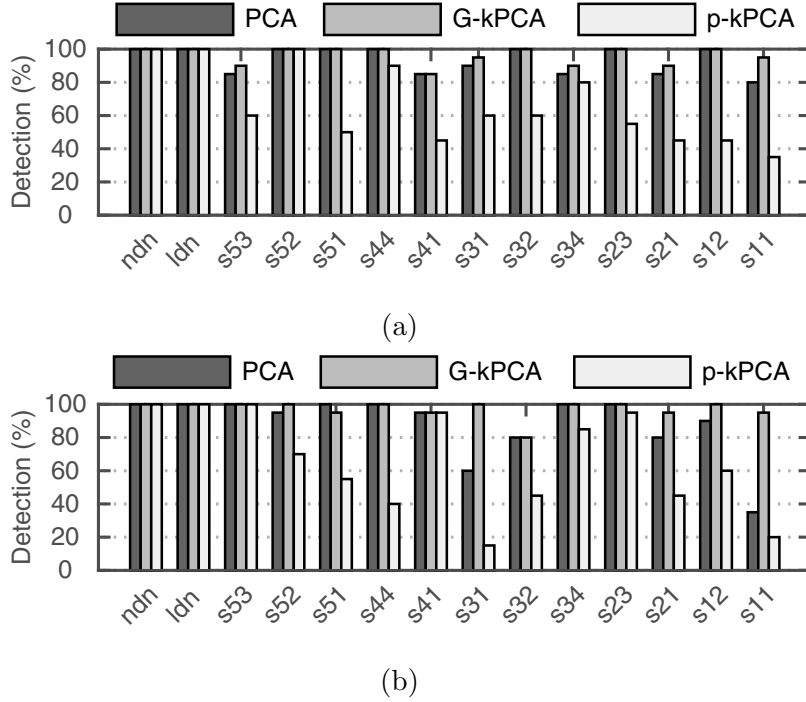
Figure 4.13: Scenario-wise detection percentages of experiments LOS (top) and N-LOS (bottom).

4.3.3.2 Scenario-wise detection performance

Here the performance of each technique is compared under true detection percentages for each scenario. Figure 4.13 provides the detection comparison of each technique scenarios in experiments LOS (top) and N-LOS (bottom). 14 scenarios are considered, two dynamic (*ndn, ldn*) and 12 static as illustrated in Figure 4.12. Highest detection percentages in general are for the two dynamic scenarios and for positions near the link such as *s12, s23,s32, s43, s52* and *s53*. Lowest detections are generally for positions far away from the link such as *s11, s21* and *s41*.

This behaviour can be explained as follows. When a person is closer to the LoS, more paths get affected. The number of paths disturbed by the person

(a) PCA
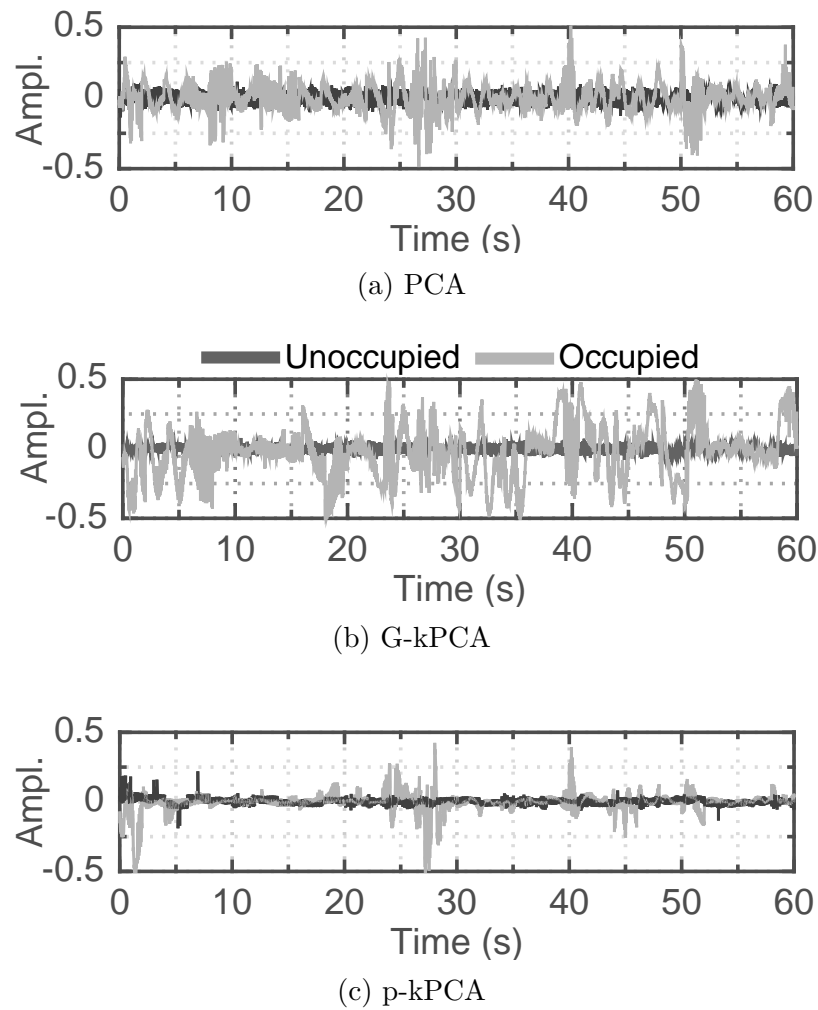
(b) G-kPCA

(c) p-kPCA

Figure 4.14: Comparison of temporal variation in *PC 2* of PCA, G-kPCA and p-kPCA for position *s32* of experiment LOS.
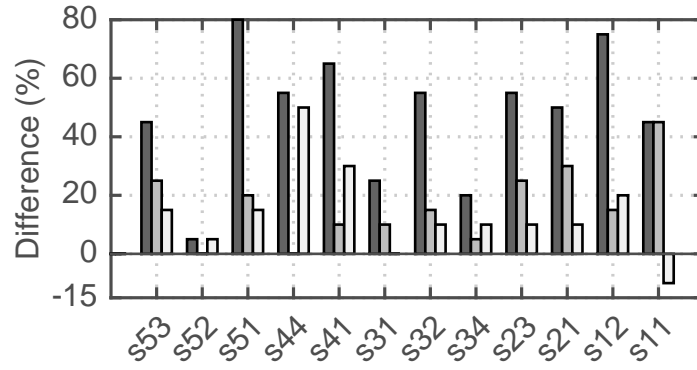
(a) PC 1
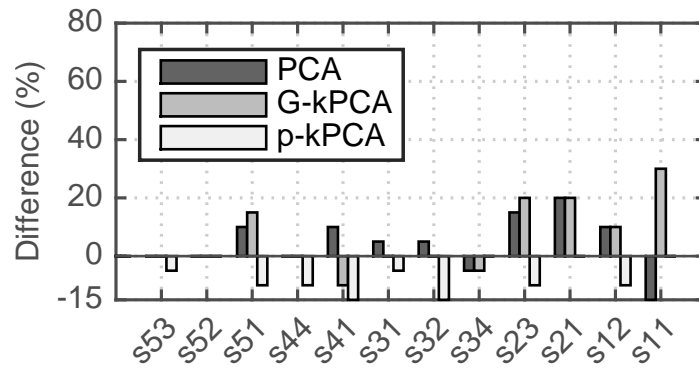
(b) PC 3
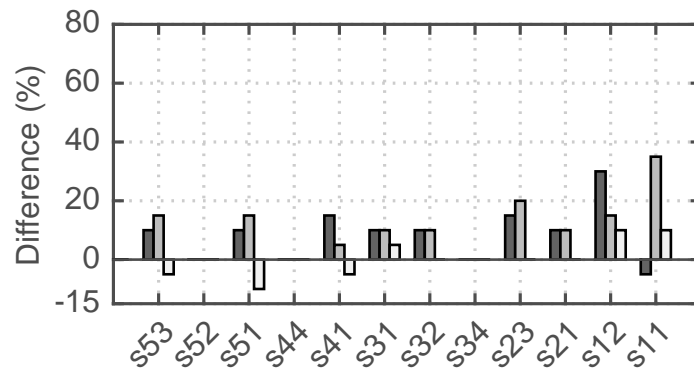
(c) PC 4

Figure 4.15: Detection percentage difference of *PC 1* , *PC 3* and *PC 4* compared to *PC 2*.

far away from the link are less and have weak amplitudes so they do not cause large perturbations in CSI. For the two dynamic scenarios performed in the experiments, the three techniques have ideal detection rates. This highlights the fact that the selected approaches are capable of detecting human motion.

4.3.3.3   Performance for a static person

Here, the detection accuracy is evaluated for a static person. The two plots in Figure 4.13 depict that Gaussian kernel PCA has a consistent detection percentage in the static scenarios. Especially for the non LoS link in experiment N-LOS, the detection performance is quite above the other two. This can be attributed to the dynamic parameter adjustment for the Gaussian kernel PCA which was explained in detail in Section 4.2.3.1. This can be further verified through the analysis of $PC2$ (which is y_i where $i = 2$ as calculated in Section 4.2.2). As an example, the behaviour of $PC\ 2$ for a static scenario, especially, for position $s32$ is illustrated in Figure 4.14.

In Figure 4.14 $PC\ 2$ of the Gaussian and Polynomial kernel PCA are compared to $PC\ 2$ of PCA. Additionally, each figure contains $PC\ 2$ of the unoccupied room as a reference. Since the three plots have different scales due to the use of different kernels for transformation, they are normalised by the maximum range of the occupied room's respective PC. The figures show that Gaussian kernel PCA has enhanced signal amplitude, whereas in the other two cases, the signal is immersed in noise. The reason is the suppression of noise through dynamic adjustment of σ in the Gaussian kernel PCA. This provides a consistently high amplitude during the presence of a person. Polynomial kernel PCA produces high amplitudes when there is only significant motion. For example in the 20 s and 30 s interval there is a sudden rise in variation. However, because the comparison is predominantly a static person in this situation, there can be only occasional movement that can cause large signal variations. For this reason, the Gaussian kernel performs better than the other two in mostly static situations. The Polynomial kernel has the lowest detection percentage which signifies that it requires further tweaking of its parameters.
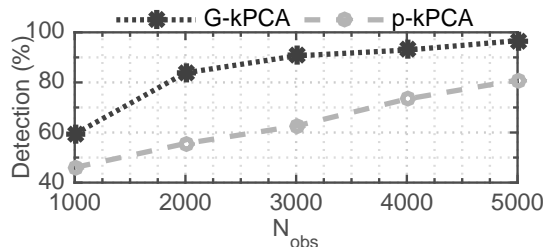
Figure 4.16: Impact of observation interval N_{obs} on detection.

4.3.3.4 Impact of the *PC* on detection

For experiment LOS, *PC 2* was selected to detect a person. Here, it is shown that the difference in detection percentages in the other three principal components compared to that of *PC 2*. Figure 4.15 shows the detection percentage difference of the other three principal components *1,3* and *4* compared to *PC 2*. *PC 1* has the highest difference and *PC 3* has the lowest for Gaussian kernel PCA and PCA. Although *PC 1* captures the highest variance fundamentally, under-performance of *PC 1* in terms of detection is due to the capture of noise in the unoccupied environment as explained in Section 4.2.4.

4.3.3.5 Impact of the observation interval on detection

In the presence detection scheme, $N_{obs} = 3000$ is selected which consists of $3\,\mathrm{s}$ intervals as the number of observations to perform dimensionality reduction and human detection. Here the impact of N_{obs} on presence detection using Gaussian kernel PCA and Polynomial kernel PCA is evaluated. Figure 4.16 illustrates the detection percentage variation with increase of N_{obs} for a single scenario in experiment LOS.

Summarizing the results, it can be stated that human detection based on Gaussian kernel PCA is quite efficient in adverse conditions such as stationary users and non LoS links. Although the Polynomial kernel PCA based approach detects human motion, it requires further tweaking of parameters.
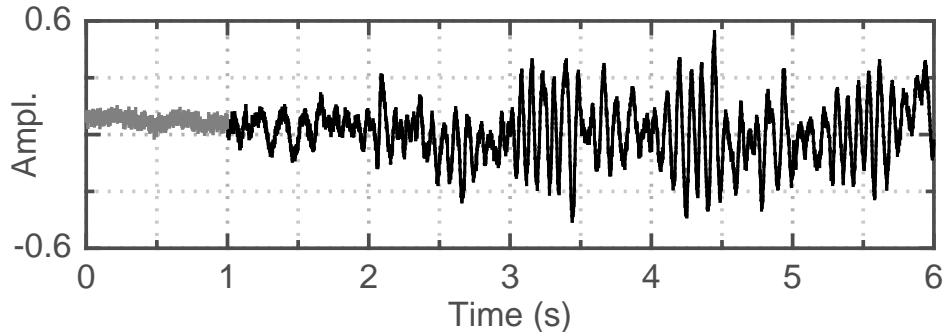
Figure 4.17: Temporal variation of *PC 2* of a walking person in Section 3.3.2 after applying G-kPCA to reduce dimensions.

4.4 Discussion

This section discusses limitations and possible augmentations to the dimensionality reduction techniques and the human presence detection scheme. This work explored only the efficient usage of CSI amplitude on human presence detection through non-linear techniques. However, the phase response can also be incorporated in the detection, especially given that the correlation between amplitude and phase are non-linear. This presence detection approach requires prior knowledge of the environment to determine principal components, thresholds and to adjust parameters. Since non-linear dimensionality reduction techniques are used, the computational complexity increases with the observation interval, N_{obs}. However, the advantage of these techniques is that the complexity is not dependent on the number of dimensions. Here, the detection performance of the scheme is not analyzed in the presence of multiple occupants. However, in such a situation a high detection performance can be expected from this system compared to a single person because multiple persons create more movements and disrupt more paths than a single person.

With enhanced detection accuracy provided especially by the Gaussian kernel, this non-linear technique can be extended to multiple occupant detection and activity recognition. As an example for an activity recognition scenario, in Figure 4.17 the second PC of the experiment performed to validate the human presence detection model in Section 3.3 is plotted. During this experiment, the person's trajectory lies perpendicular to the link. The person's movement is plotted during

the interval $1\,\mathrm{s}$ to $6\,\mathrm{s}$, as a reference, occupied room noise is plotted during the interval $0\,\mathrm{s}$ to $1\,\mathrm{s}$.

4.5 Conclusions

In this work, a nonlinear approach to detect human presence using CSI provided by off-the-shelf WiFi chipsets is provided. First the effect of humans on CSI amplitude is analyzed using an experiment and using that the linear/non-linear relationships among CSI subcarrier amplitudes are discovered. These correlations are exploited by introducing non-linear techniques to reduce CSI dimensions and then human presence is detected with just two WiFi transceivers. The overall detection performance of the system is evaluated through true and false positives. It is shown that especially the Gaussian kernel can achieve a higher detection performance than the linear approach.

Chapter 5

Fall Detection Architecture

In this chapter, an ambient fall detection approach based on WiFi CSI, termed FallDeFi, is presented. As highlighted in Section 2.2.2 prior work on fall detection using CSI have only considered time domain approaches for fall detection and time-frequency analysis is used for activity recognition only. In this work, the latter approach is followed because time-frequency feature extraction and analysis has significant advantages compared to pure time domain or frequency domain analysis. The main advantage is the amplitude information in both time and frequency axes are available for analysis which has enabled researchers to identify the Doppler shift in WiFi signals.

Even though existing WiFi activity recognition systems classify falls among other generic activities, these systems do not analyze the robustness of fall detection among activities that have similar characteristics to falls. As an example, daily activities like sitting down, standing up, jumping, and bending and picking up objects have similar time-frequency patterns that can mislead the system. Therefore, specialized features that distinguish these activities are required especially given that fall detection is closely related to safety and health of elderly people. Additionally, these features have to be resilient to changes in the environment enabling easy deployment of a pre-trained system and do not require retraining due to changes in the environment. In FallDeFi, fall detection of commodity WiFi devices is combined with established signal processing and feature extraction techniques from radar based fall detection to create a low cost, reliable and environment resilient fall detection approach.
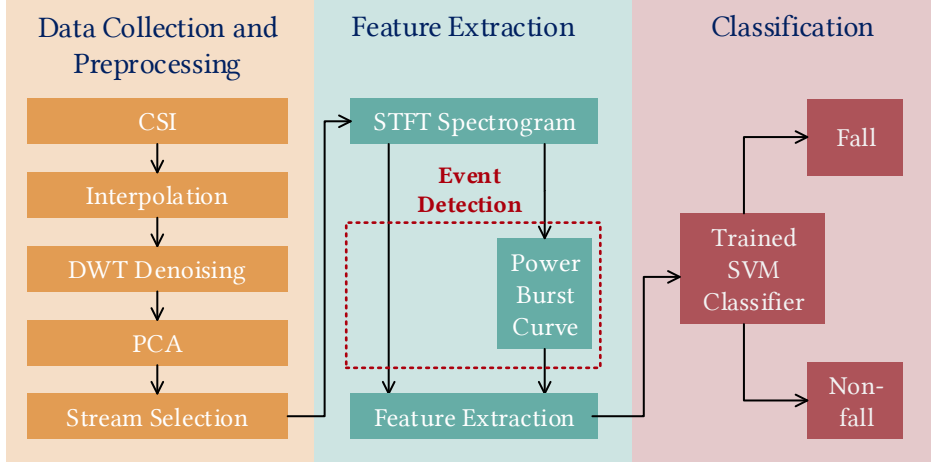
Figure 5.1: FallDeFi system architecture.

5.1 System Overview

The proposed system, FallDeFi, is an ambient/device-free, indoor fall detection system that uses commodity WiFi devices as the physical fall monitoring infrastructure. FallDeFi continuously monitors the occupant activities by processing CSI captured from WiFi signal receivers to identify falls. When it detects an activity, it classifies this as either a fall or a non-fall using a pre-trained model. However, this model has to be accurate and resilient to variations in the channel caused by changes in the monitored environment either as a result of changes in furniture locations, other big items in the space or even WiFi transceiver/antenna positions. To address these requirements, the system, illustrated in Figure 5.1 is built in a modularized manner: (i) Data Collection and Preprocessing, (ii) Feature Extraction and (iii) Fall Classification, inspired by radar fall detection [12] and WiFi activity recognition [26] architectures. In this section, the objectives of each module and the components therein are summarized while details about the motivations for selection and the related state of the art of each module and the components are mentioned under the heading of each module in Section 5.3.

The Data Collection and Preprocessing module is responsible for collecting the raw CSI signals from WiFi devices and processing them to obtain a clean and informative spectrogram for time-frequency analysis. To this end, this module addresses three issues that occur with CSI : (i) obtaining a fixed rate of CSI amid

packet losses due to weak signals in through-wall and non-LoS links, (ii) removing noise in the subcarriers without losing high frequency components especially when the signal is weak, and (iii) extracting the most effective streams from the subcarriers as input to the spectrogram for increased computational efficiency. To solve these issues linear interpolation, Discrete Wavelet Transform (DWT) based noise filtering and PCA based stream decorrelation and selection are used respectively.

The Feature Extraction module has three main functions: (i) detection of start and end of an event, this component should detect 100% of human induced events, (ii) extraction of features during the detected event interval, and (iii) selection of features that classify events. Since a spectrogram is used to extract features, events are detected from a Power burst Curve (PBC) which is commonly used in radar fall detection to detect events in a spectrogram. Once an event is detected, features are extracted both from the spectrogram and the PBC. While the spectrogram provides spectral features, the PBC provides time and amplitude related features. Once the features are extracted, sequential forward selection is used to determine the environment independent features and feed them to the classifier. Finally, the Fall Classification module classifies the detected events into falls and non-falls using the extracted features. Algorithm 1 provides the detailed flow of events and the inputs and outputs of each module from the moment CSI is collected until falls are detected, as implemented in FallDeFi.

5.2 CSI Preliminaries

WiFi devices that operate in IEEE 802.11a/g/n/ac use OFDM as the modulation scheme to send data using multiple subcarriers in a single 20 MHz WiFi channel. The receiver measures a discrete CFR in time and frequency as phase and amplitude in the form of CSI for an antenna pair. The CSI amplitude and phase $H(t; f_i)$ of a subcarrier i at time t and frequency f_i can be modeled as a combination of *unaffected* and *affected* components due to human motion:

$$H(t, f_i) = H_u(t, f_i) + \sum_{k \in P_d} a_k(t) e^{j2\pi \int_{-\infty}^{t} f_{D_k}(v)dv} e^{-j\kappa(t,f_i)} \tag{5.1}$$

Algorithm 2: FallDeFi, fall detection algorithm

  input : CSI_{raw} - CSI measurements of a t_{int} second interval from
     $m \times n$ antenna pairs
  output: Fall detection

12 **Initialization :** $\hat{N}_{th}, PBC_{th}$;

13 **foreach** *antenna pair* **do**

14 **foreach** *subcarrier* **do**

15 $CSI_{interp} \leftarrow$ LinearInterpolate (CSI_{raw});

16 $CSI_{wden} \leftarrow$ WaveletDenoise (CSI_{interp});

17 **end**

18 PCs $\leftarrow$ PCA (CSI_{wden});

19 selected PCs $\leftarrow$ SelectOptimumPrincipalComponents (PCs);

20 **foreach** *selected PC* **do** S_{raw}(n,k) $\leftarrow$ ComputeSpectrogram
   (selected PC);

21 **end**

22 S_{avg}(n,k) $\leftarrow$ AverageAllSpectrograms $(S_{raw}$(n,k)$)$;

23 S(n,k) $\leftarrow$ ProcessSpectrogram $(S_{avg}$(n,k)$)$;

24 PBC(n) $\leftarrow$ ConstructPowerBurstCurve (S(n,k)) ;

25 **if** $PBC(n) > PBC_{th}$ **then**   // Fall-like event detection

26 feature vector $\leftarrow$ ExtractFeatures (PBC(n), S(n,k));

27 Fall $\leftarrow$ SVM(trained model, feature vector) // Classify falls
  and non-falls

28 **end**

---

where $H_u(t, f_i)$ is the component unaffected by human motion and $a_{k(t)}$ is the amplitude of a single path scattered of a human, f_{D_k} is the Doppler frequency shift due to human motion and $e^{-j\kappa(t,f_i)}$ is the unknown phase shift due to carrier frequency offset and packet detection delay.

PCA captures a measure of Doppler shift as the varying rate of the propagation path length in the time domain [26]. This signal can be converted to the time-frequency domain using STFT to extract features. Due to the unknown phase shift, the CSI phase is unusable without calibration. Therefore, activity recognition schemes eliminate this by obtaining only the amplitude response, $||H(t, f_i)||$ [26; 101; 102; 120]. However, using the amplitude response alone loses the direction of the Doppler shift.

5.3 CSI Signal Processing

Due to weak signals in certain links caused by non line of sight connections and wall penetration in indoor environments, some packet losses can occur ($< 1\%$). To counter the loss of information due to packet loss, two steps are taken,

a) packet transmission rate is set well above the Nyquist sampling rate. In this work, the transmission rate is set as ($1000 \, \text{Hz}$) well above the rates usually required for a successful fall detection (the Nyquist sampling frequency of a fall is $< 350 \, \text{Hz}$ as in eq. 5.6 for WiFi at $5.2 \, \text{GHz}$ and even lower at $2.4 \, \text{GHz}$).

b) signal strengths are interpolated on individual subcarriers. In order to obtain the same sample rate for all the traces, linear interpolation is preferred.

As mentioned in Section 3.2.2.1, Intel 5300 chipsets extract CSI from WiFi receivers. CSI subcarriers from these devices consist of noise from several sources e.g. radio frequency interference, multipath propagation and noise from imperfections of the chipsets. When subcarrier streams are decorrelated through PCA, noise present in the subcarriers makes the first PC unusable. Therefore, some research work has neglected the first PC and relied only on the information in other principal components [26]. However, as the first PC contains the most information as it corresponds to the highest eigenvalue among all the principal components, neglecting this could result in losing information. First, in Section 5.3.1 a noise filtering approach is introduced and then in Section 5.3.2 its impact on the first PC is discussed.

5.3.1 Noise Filtering

Initially, several time and frequency domain noise filtering approaches were considered. Time domain approaches such as median or mean filters can distort the signal, and vital high frequency components may also be lost. Although frequency domain approaches like bandpass filters (e.g. Butterworth) remove the noise well for narrow bands, this is an out-of-band noise filtering technique. If a bandpass filter with a larger passband is used, the residual noise will have a higher contribution and a smaller passband cuts off vital signal components. Since CSI has

noise in all bands, Discrete Wavelet Transform (DWT) which is an in-band noise filtering technique is chosen for this work. Through careful parameter selection, DWT filtering eliminates in-band noise, preserves high frequency components and introduces less distortions to the signal.

In DWT-based noise filtering, by transformation into the wavelet domain the signal is divided into multiple frequency levels called *wavelet levels* (L) that consist of *approximate* α^L and *detailed* coefficients β^l. The denoising procedure involves decomposing the original signal x_n,

$$\alpha_k^L = \sum_{n \in \mathbb{Z}} x_n g_{n-2^L k}^L \tag{5.2}$$

$$\beta_k^l = \sum_{n \in \mathbb{Z}} x_n h_{n-2^l k}^l, l \in \{1, 2, \ldots, L\} \tag{5.3}$$

threshold detailed coefficients β_k^l and reconstruct the synthesized signal.

$$x_n = \sum_{k \in \mathbb{Z}} \alpha_k^L g_{n-2^L k} + \sum_{l=1}^{L} \sum_{k \in \mathbb{Z}} \beta_k^l h_{n-2^l k^l} \tag{5.4}$$

where the gs and hs are the wavelet basis functions (in this work symlet basis functions are selected). The highest wavelet level, L, is considered noise, based on that, noise level and the threshold for that level is estimated. Subsequently, the threshold is adapted to lower wavelet levels based on signal variance at that level, and then noise is cut off in lower wavelet levels without distorting the signal component. The underlying model of the noisy signal β_k^l is of the following form

$$\beta_k^l = b_k + \sigma e_k \tag{5.5}$$

where b_k is the denoised signal, e_k is the noise level and σ is the noise variance.

As an example, for a Nyquist sampling rate of $1000\,\mathrm{Hz}$, assuming that human falls do not occur beyond $250\,\mathrm{Hz}$, the highest wavelet level $250\,\mathrm{Hz}$ to $500\,\mathrm{Hz}$ is considered noise. Since the maximum velocity of a falling human just before hitting the ground is typically $5\,\mathrm{ms}^{-1}$ [121], applying the Doppler frequency formula

,

$$\delta f_{max} = \frac{2v}{c} f_c \tag{5.6}$$

to a 5.2 GHz WiFi signal returns a maximum Doppler frequency of $\delta f_{max} \approx$ 175 Hz, which is well below 250 Hz. Thus all frequency components above 250Hz can be removed, however, to avoid distortions in other frequencies, not all high-frequencies components are removed but only those with limited influence on the signal. This is illustrated in Figure 5.2. The parameters that are initialized for wavelet denoising are discussed in Section 5.6 under *System calibration*.

Figures 5.3a and 5.3b illustrate the transformation of a noisy subcarrier amplitude to a noiseless signal through the selected strategies. The effect of noise filtering on the first principal component is illustrated in Figure 5.4. The effect of this noise filtering approach on fall classification accuracy is illustrated in Section 5.7.1.1.

5.3.2   Subcarrier Decorrelation and Stream Selection

A single antenna pair provides 30 streams of CSI from 30 subcarriers and they are mostly correlated, especially when there is an activity [26]. However, for time-frequency analysis, as few streams as possible are required to reduce the computational overhead in constructing the spectrogram for feature extraction. Methods like averaging a set of streams or selecting a subset of streams do not capture the variations in all the subcarriers effectively. Therefore, the CSI streams are decorrelated using PCA and the minimum number of resulting orthogonal streams, i.e, the principle components that capture $x\%$ of the variance contained in the 30 subcarriers are selected.

In PCA, a linear transformation ensures the projection of CSI subcarrier amplitudes to a new orthogonal coordinate system such that the highest variance is present in the first coordinate, the next highest variance is present in the second coordinate, and so on. The linear transformation in PCA ensures that the cumulative variance in the resultant coordinates is equal to the cumulative variance in all the original subcarriers. These coordinates are called the Principal Components (PCs), and the amount of variance in a principle component is explained by the respective eigenvalue obtained through PCA. Due to the DWT-based noise
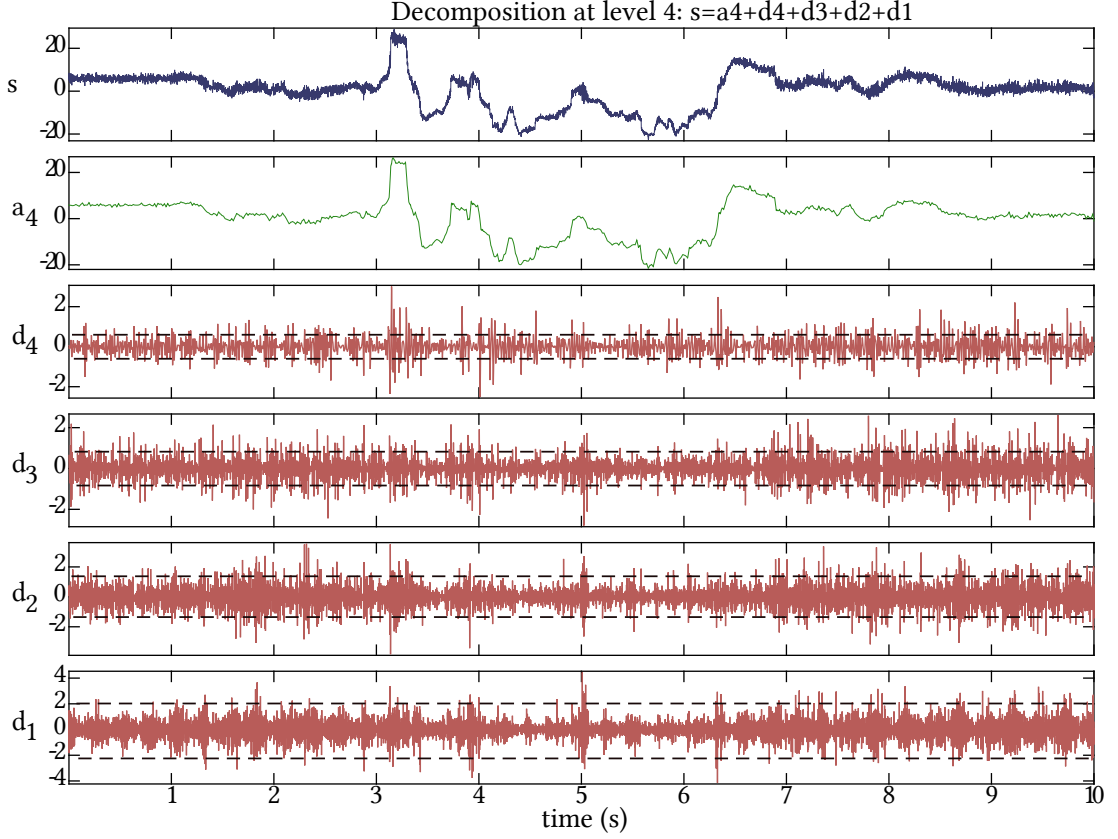
Figure 5.2: DWT filtering used in this work. s is the noisy signal and a 4 level sym3 wavelet is applied to s. d_1, d_2, d_3 and d_4 are detailed coefficients and a_4 is the approximate coefficient obtained from wavelet transform. The thresholds of detailed coefficients are marked in the dashed blue line.

filtering and PCA as per Section 5.3.1, the first two to three PCs are, on average sufficient to capture the majority of the variance in the original 30 subcarriers. The number of PCs for a particular activity are selected to satisfy the equation,

$$\text{Captured variance} = \frac{\text{Cumulative variance of selected PCs}}{\text{Cumulative variance of all 30 subcarriers}} \qquad (5.7)$$

The selection of PCs satisfying the value of *captured variance* as in eq. 5.7 is empirical and it is described in Section 5.7.1.2. As PCA is an unsupervised machine learning approach, selecting a fixed number of PCs can include noisy

components for some activities and lesser information for others. Unlike previous research that use a fixed number of PCs [26; 101], FallDeFi selects the required minimum number of PCs dynamically for each activity depending on the number of PCs that satisfy the predefined captured variance.

Figures 5.3 and 5.4 show how the noise filtering at subcarrier level has a higher impact on the noise level of the first principal component than on the subcarrier. As highlighted in Section 5.3.1, the figure shows that the noise present in raw subcarriers can result in high frequency noise in the first principal component, especially on occasions when there are no activities. Additionally, it can be observed that for noisy links, high frequency noise has a higher impact on the first principal component compared to the signal itself. Figure 5.5 shows the first three principal components for the same activity when de-noising is applied before PCA.

(a) Subcarrier w/o denoising.   (b) Subcarrier after Wavelet denoising.
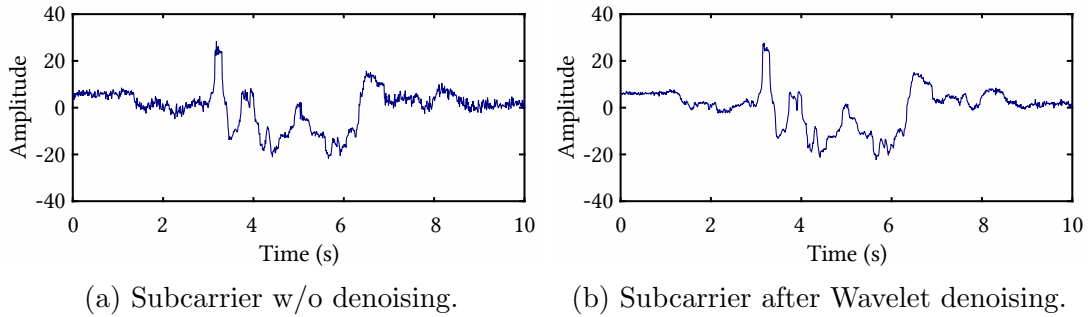
Figure 5.3: Comparison of the amplitude variation of 1st subcarrier before and after wavelet denoising

(a) PC 1 before denoising.   (b) PC 1 after noise filtering.
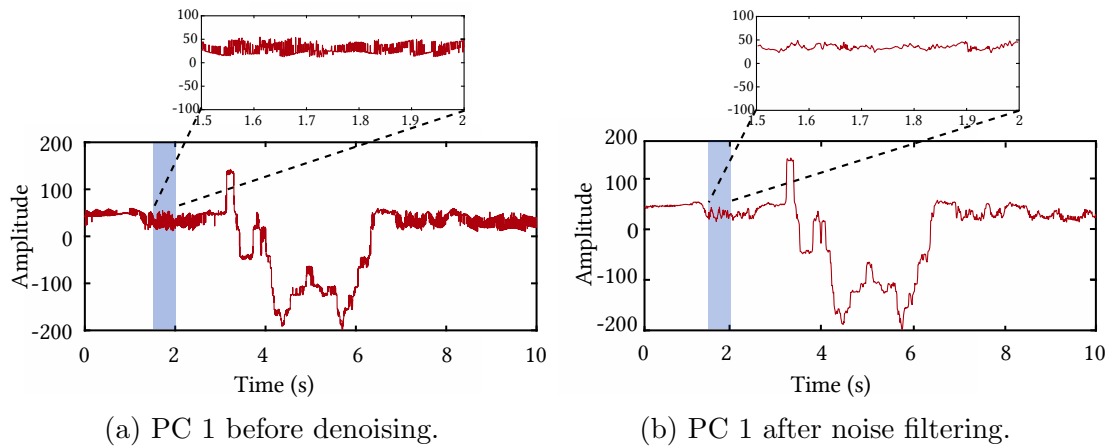
Figure 5.4: Effects of de-noising are more prominent in PC 1 of Fig. 5.4b than in individual subcarriers in Fig. 5.3b.

(a)   (b)   (c)
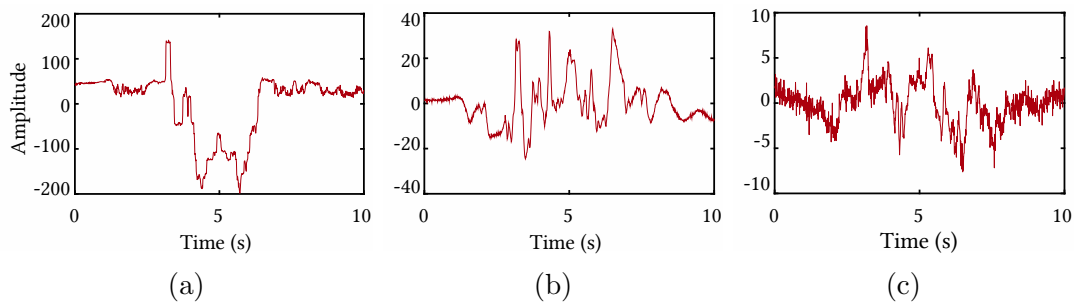
Figure 5.5: Comparison of the resultant first three PCs after wavelet de-noising and then PCA.

5.4 Time Frequency Analysis

5.4.1 Spectrogram and its Parameter Selection

After CSI data collection and preprocessing, a transformation to time-frequency domain is necessary to perform the feature extraction. For time-frequency analysis, various linear and non-linear techniques exist especially among radar fall detection [12] approaches. Among these, non-linear methods tend to distort the frequency components generated by falls. Additionally, among the various time-frequency feature extraction methods, STFT spectrogram features have outperformed DWT features in recent radar fall detection research [77]. Therefore, STFT is selected for time frequency analysis.

In STFT, the frequency resolution is inversely proportional to the time resolution. Hence, this requires to select an optimal window size to obtain satisfactory time and frequency resolutions for the required application. With CSI measurements for human falls having frequencies up to a theoretical maximum of $175\,\mathrm{Hz}$ in the $5\,\mathrm{GHz}$ band, and as changes in time occur in the sub-second range, an FFT window size of 512 samples at a sampling rate of $1000\,\mathrm{pkts/s}$ is selected. The overlap size of two windows is chosen to be 256 samples for two reasons: *a)* it is the optimum overlap size for the Hamming window used in FFT, *b)* higher value increases the computational efforts as it introduces high interpolation. The selected parameters provide a frequency resolution of

$$\frac{sample\ rate}{FFT size} \approx\ 2\,\mathrm{Hz} \tag{5.8}$$

and a time resolution of

$$\frac{window - overlap}{sample\ rate} \approx\ 0.1\,\mathrm{s} \tag{5.9}$$

. Taking into consideration that a fall can occur within a minimum time duration of 2s and a minimum frequency range of 30Hz, the selected STFT parameters can collect a minimum of 20 data points in time axis and 15 data points in frequency axis for feature extraction.

First, the optimum number of principal components that capture $95\,\%$ of

Figure 5.6: Comparison of spectrograms for falls and fall like activities. (a) Trip, (b) Slip, (c) Lose balance, (d) Lose consciousness, (e) Sit down on the floor, (f) Stand up from the floor (g) Sit down on a chair, (h) Stand up from a chair, (i) Bend and pick up an object, (j) Jump, (k) Walk and (l) No event

Figure 5.7: (a) Raw spectrogram image, (b) Binary image of the raw spectrogram, (b) Binary image after Morphological processing, (d) Binary image after disconnected region removal

total variance of all the subcarriers is obtained. Then STFT is applied on the selected principal components individually and the resultant spectrogram matrices averaged to compute a mean spectrogram matrix for each activity according to [26; 101]. As the principal components are orthogonal to each other and each of them consist of unique frequency components, by averaging the spectrograms, a spectrogram with a higher amount of information can be obtained. The spectrograms obtained through this process for four types of falls (trip, slip, lose balance and lose consciousness) and other activities (sit, stand, bend and pick an object, jump, walk and no event) are illustrated in figure 5.6.

5.4.2 Spectrogram Processing

Before extracting the features corresponding to falls and fall-like activities, the main event has to be isolated to obtain a clear spectrogram that characterizes only the falls and fall-like events. Therefore, time-frequency energies not occurring within the main event interval are treated as noise and are removed from the spectrogram.

In the literature, spectrogram thresholding is a common technique that segments spectrograms [78; 79]. To determine the threshold, the noise at high frequencies of the spectrogram is estimated. When the histogram of the amplitudes at high frequencies ($>250\,\mathrm{Hz}$) is constructed, it can be observed that it can be approximated with a Gaussian distribution which is also consistent with previous time-frequency analysis work [122]. This is due to the fact that events do not occur beyond $175\,\mathrm{Hz}$ and this region is dominated by noisy parts. If low frequency amplitudes were also considered, the histogram would deviate from the Gaussian distribution and would become more skewed, due to the effects of human movement events at low frequencies. Hence, only the frequencies beyond $250\,\mathrm{Hz}$ are chosen to estimate the spectrogram segmentation threshold $\hat{N}_{th}$. The spectrogram segmentation threshold can be defined as

$$\hat{N}_{th} = \mu + k\sigma \quad k\epsilon R \tag{5.10}$$

where μ is the mean and σ is the standard deviation of the Gaussian approxi-

mation of the noise histogram from $250\,\mathrm{Hz}$ upwards. The value k is selected to tune the spectrogram segmentation threshold for different environments so that the weak energy parts around the extreme frequencies of the main event are preserved and noise energy is segmented (Figure 5.7). An average value for k can be determined by measuring noise in several spectrograms for a particular link.

$$S(n,k) = \begin{cases} S(n,k) & \text{if } S(n,k) \geq \hat{N}_{th} \\ 0 & \text{if } S(n,k) < \hat{N}_{th} \end{cases} \tag{5.11}$$

After spectrogram segmentation, disconnected regions may appear due to weak reflections off the limbs or other minor motions during a fall, which can cause weak energy components to be scattered around the extreme frequencies of the main event. For this reason, morphological dilation is performed [123] to connect close yet broken spectrogram segments that are within a predefined frequency. The regions that were not connected are then removed from the main event region, specified as disconnected region removal. This procedure is illustrated in Figure 5.7.

5.4.3 Event Detection

A fall event occurs typically within a 2 to $3\,\mathrm{s}$ period [100; 121], comprising high frequency/energy components that are not so apparent in other activities which can also be accompanied by low frequency/energy events prior to and after a fall. Therefore, it is important to localize the fall from other events to feed the fall specific features to simplify the fall classification. Thus, the event detection step should be able to effectively localize only the high energy/frequency fall and fall-like activities. For these reasons, the typical time domain event detection methods cannot be used. Instead, a PBC [77] is used to detect and localize high frequency events. Unlike Doppler radars, the current spectrogram $S(n,k)$ does not associate negative frequencies as the motion direction is not captured by CSI. Therefore, the PBC is only defined for positive frequencies

$$PBC(n) = \sum_{k=k_l}^{k_u} |S(n,k)|, n = 1, 2, \ldots, N \tag{5.12}$$

where k_l is the lower frequency bound and k_u is the upper frequency bound. Even though a fall may have frequency components in excess of $100\,\text{Hz}$ in the $5.2\,\text{GHz}$ WiFi band, the strongest signal energy of high frequency events including falling and other similar activities are concentrated between $5\,\text{Hz}$ to $30\,\text{Hz}$. The frequency components within this range correspond to reflections from a large body of mass like the torso. Therefore, $k_l = 5\,\text{Hz}$ and $k_u = 25\,\text{Hz}$ are set, signal powers are summed between this range and when this exceeds a predefined threshold, PBC_{th}, it is considered as a fall-like event. The threshold, PBC_{th} is determined as

$$PBC_{th} = \sum_{k=k_l}^{k_u} \hat{N}_{th} \qquad (5.13)$$

where $\hat{N}_{th}$ is the estimated noise threshold in eq. 5.10, which is dependent on the link. Note that the upper frequency of the PBC is 25Hz which is set 5Hz lower than 30Hz to provide a margin for detecting low energy signals within the 5-30Hz frequency range. In this case, the power burst curve threshold PBC_{th} is determined by summing the noise amplitude, $\hat{N}_{th}$, estimated in previous section within the 5 to 25 Hz frequency range. This process is illustrated in Figures 5.8a and 5.8b. Unlike time domain prescreeners, the main advantage of this approach is the isolation and detection of events within a specific frequency range.

5.5   Feature Extraction and Classification

This section explains the set of features that were extracted from the spectrogram for classification. From previous research that involve time-frequency features, a pool of features that have been successful were chosen and tested on the traces collected for both falls and non-falls. First, from the collected traces, the features that correctly classify the falls were identified, and termed as *Original Features (OFs)*. From these features, a feature subset which is more robust when environment changes occur is chosen and is called *Selected Features (SFs)*. To select these two types of features, a commonly used yet simple sequential forward selection algorithm [124] is used. This algorithm can be summarized as follows:

i) Start with an empty feature set,
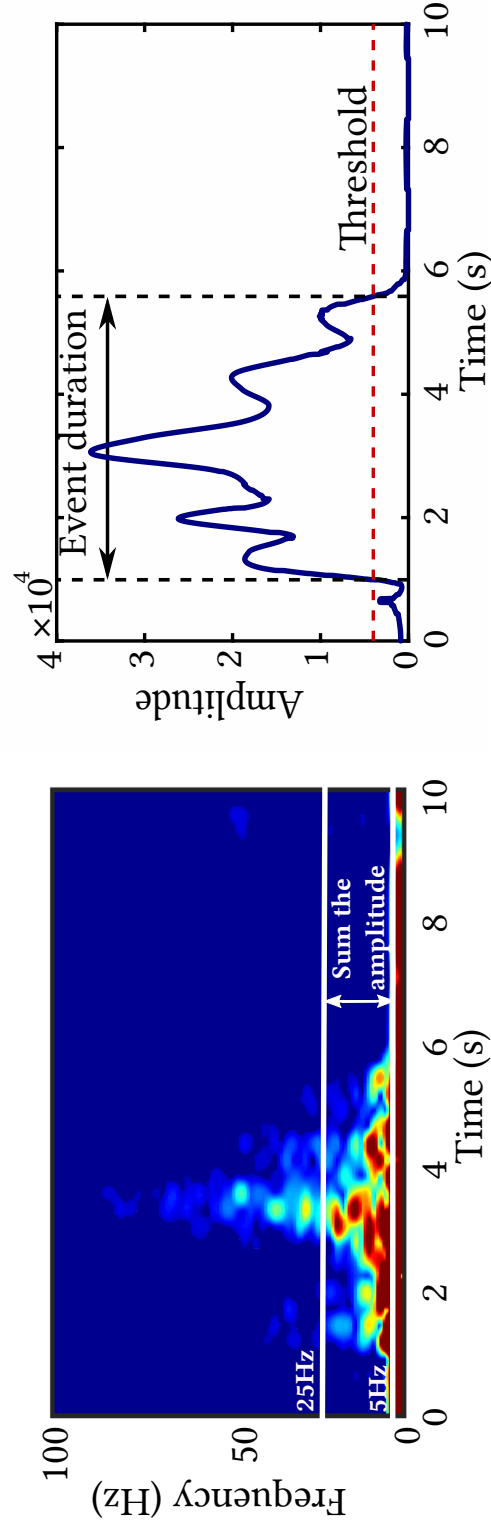
(a) The spectrogram of a composite event that also consists a fall.

(b) Power burst curve of the event in Figure 5.8a.

Figure 5.8: Creation of a power burst curve (b) from a spectrogram (a). The threshold in Figure 5.8acan be adjusted to isolate high frequency or low frequency events

ii) Determine the next feature that produces the best classification performance when used collectively with previously chosen feature/features,

iii) Update the best feature set,

iv) Repeat from step 2 until all the features are tested.

Two criteria are used to select *OFs* and *SFs* in this algorithm. For *OFs*, the criterion is the highest accuracy when trained and tested on the same data. For *SFs*, the criterion is highest accuracy when trained and tested on two different data sets. Table 5.1 illustrates both sets of features and the classification results of these two features are explained in Section 5.7.2.

5.5.1 Original Features

As shown in Table 5.1, the extracted *spectral features* are as follows.

(i) *Extreme frequency features* [77; 123]: mean, standard deviation and maximum of the extreme frequency curve.

(ii) *Torso frequency features*: mean, standard deviation and maximum of the torso frequency curve. The percentile method is used to obtain the torso frequency curve which is the least affected method by noise and other limb motions according to [79].

Table 5.1: Extracted features from sequential forward selection

	Spectral features	Power burst curve features
Original features	extreme frequency features, torso frequency features, ratio of max. extreme and max. torso freqs, spectral entropy, fractal dimension	event duration, ratio of energy above and below the PBC threshold
Selected features	spectral entropy, fractal dimension	event duration

(iii) The ratio between maximum extreme frequency and maximum extreme torso frequency. Falls have both higher extreme frequencies and higher torso frequencies.

(iv) *Spectral entropy.* This is a normalized feature and measures the textural properties of a fall (randomness in the distribution of energy in a spectrogram) [125]. As an example, a fall has a higher entropy than other events at both lower and higher frequencies due to a high fluctuation in amplitudes. The spectral entropy, H is computed as:

$$H = -\sum_{i=k_l}^{k_u} p(n_i) \ln p(n_i) \tag{5.14}$$

where k_l and k_u are lower and upper frequency bounds. $p(n_i)$ is the normalized power spectral density

$$p(n_i) = \frac{\hat{P}(k, n_i)}{\sum \hat{P}(k, n_i)} \tag{5.15}$$

where

$$\hat{P}(k, n) = \frac{1}{N} |S(n, k)|^2 \tag{5.16}$$

of spectrogram amplitudes $S(n, k)$ mentioned in Section 5.4.2. Even though a range of frequencies were picked initially, the frequency ranges that produced the best fall classification from the feature selection algorithm are the following: (a) 1 to 10 Hz, (b) 10 to 30 Hz and (c) 30 Hz to max frequency.

(v) *Fractal dimension*: This feature, calculated using the Hausdorff dimension [126] method, measures the roughness of the extreme frequency curve during the detected event interval. Since a fall includes whole body movements, movements including hands, body and head and leg unlike ADLs that include only certain parts of the body, it can be observed that the extreme frequency curve has a chaotic nature. Additionally, this is a normalized feature which can extract features robust to environment changes.

The extracted *power burst curve* features are as follows:

i) *Event duration*: Typically a fall has a duration of 2 to 3 s [100; 121], whereas other events have a range of durations depending on the activity. This is a characteristic unique to a fall which is independent of the environment and additionally, it can distinguish other events that have high frequency components yet the duration is either longer or shorter in length.

ii) *Ratio of energy above and below the PBC threshold*: Some events have high frequencies but low energy in those frequencies, however, falls have high powers in high frequencies as well as a significant amount of energy in those frequencies. This helps to separate falls from other such activities.

5.5.2 Selected Features

One of the major goals of the fall detection scheme is to classify falls with high robustness to environment changes. As shown in Table 5.1, the feature subset that best satisfied this criterion are: *event duration, spectral entropy* (1 to 10 Hz & 10 to 30 Hz) and *fractal dimension*. Note that spectral entropies above 30 Hz are not being used for classification. For positions that are far from the link, especially when the link is weak due to non line of sight conditions, even when high frequency components are produced during falls, they are not captured by the system. As it is visible in Figure 5.6, low frequency components have higher energy than higher frequency components. When falls occur far from the link, amplitudes of higher frequencies are further weakened. Therefore, such falls tend to get misclassified with fast sitting down or standing up of stronger links. However, fractal dimension of the extreme frequency curve still produced a robust classification during environment changes. The classification results of the selected features compared to originally extracted features are presented in Section 5.7.2.2.

5.5.3 Classifier Selection

Binary classification is performed to separate a fall from other activities because of two reasons. Firstly, only a fall requires to be distinguished from a non-falling event. Secondly, the features are specially suitable for distinguishing falls from other events and not the individual activities. As performance of a classifier is
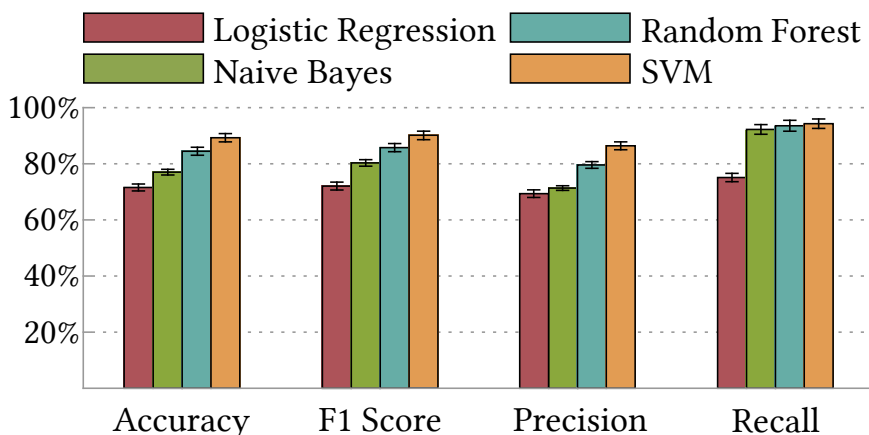
Figure 5.9: Classifier comparison. Performance of SVM is higher than others in all the metrics.

mostly dependent on the data set, several classifiers (logistic regression, naive Bayes, Support Vector Machine (SVM) and random forest) were compared before choosing one. SVM was implemented using the LIBSVM [127] tool set, and the hyper-parameters σ and c were chosen through five-fold cross validation. For random forest, the number of trees was set to 25 as it provides high classification accuracy and avoids significant over-fitting. Then the performance of each classifier was compared as shown in Figure 5.9. Based on these results, SVM with Radial Basis Function kernel provided the best performance out of the four classifiers and it was used for the rest of the analysis.

5.6 Evaluation Setup

In this section, the environments in which the experiments were performed, the hardware and software tools, the types of data sets that were gathered, the type of activities that were conducted to distinguish falls, and finally, the evaluation metrics that were used to measure the performance of FallDeFi are presented.

5.6.1 Experimental Environments

To evaluate the performance of the fall detection system across diverse conditions, falls and other activities in several environments were performed. Figure 5.10

(a) Apartment

(b) Lab

(c) Bathroom

Figure 5.10: Experimental environments. [©David Rojas]

illustrates the experimentation environments. Marked in *yellow* crosses is the area that both falls and other activities were performed in. To avoid injuries to the volunteers due to the limited space, in the area marked in *blue* crosses, other activities are performed. As shown in the figure, these environments include an apartment, a laboratory, and a bathroom and a toilet. The experimentation area of the apartment consists of two bedrooms, a corridor and the kitchen. The experiments were conducted in typical indoor environments to encompass cluttered, uncluttered, LoS, non-LoS and through-wall scenarios. The link types of all environments are as follows:

(i) kitchen - 4 m link, LoS cluttered environment,

(ii) corridor - 9 m link, LoS uncluttered environment,

(iii) lab - 7 m link, non-LoS cluttered,

(iv) two bedrooms - 5 m link, through wall (single wall with decoupled and insulated plaster partition), and

(v) bathroom and toilet - 5 m link, through wall (two walls, each having a plaster partition).

5.6.2 Hardware and Software Setup

FallDeFi was implemented using commercial WiFi devices, the transmitter and receiver in each link consist of two Linux laptops. Each of them is equipped with two external omnidirectional antennas. In each laptop, a WiFi network interface card with the IEEE 802.11n Intel WiFi Link 5300 chipset was installed, the driver and firmware were modified in the laptops to function as transmitter and receiver. The WiFi chipsets in the laptops operate in monitor mode which is one of the IEEE 802.11n WiFi modes to sniff packets in a particular channel. CSI Tool [119] was used to analyze the data collected from the chipsets. The source code and data sets are available in [128].

Table 5.2: Data collection in two different days. Group A has five data sets for Corridor, Bedroom, Kitchen, Lab and Bathroom for falls and non-falls, while Group B has only four.

Group	Measurements	Corridor	Bedroom	Kitchen	Bathroom	Lab	Total
A	Falls	33	21	39	35	49	177
	Others	58	70	99	81	86	394
B	Falls	49	30	30	40	-	149
	Others	77	91	90	92	-	350
	Diff. in Days	7	33	33	10		
Changes from A to B	Diff. in environment	Tx. moved by .5m	+1 person	+1 person, furniture moved nLoS	+1 person, Tx moved by .5m		

5.6.3 Execution

During the data collection period for each activity, the transmitter sends 1000, 100 Byte packets over a t_{int}=10 s duration resulting in a data rate of 800 kb/s. Therefore, 10 000 packets were collected corresponding to 30×10000 CSI values for each antenna pair for each activity. All experiments were performed in the 5.2 GHz band because it has less interference than the 2.4 GHz band and the bandwidth of activities are higher than in 2.4 GHz. Table 5.2 illustrates the amount of repetitions of falls and other activities that were conducted in each environment. In each environment except the lab, two data sets A and B were collected on two different days with changes in the environment. The lab data set is used as the training environment for robustness studies in section 5.7.2.2. The table further illustrates the changes from data sets A to B. Collection of data sets A and B for bedrooms and kitchen followed a gap of 33 days while corridor and bathroom data sets had gaps of 7 and 10 days between them. Additionally, from A to B the corridor and bathroom transmitters were moved by 0.5 m from the original positions to synthetically generate environment changes as these two places are not cluttered by objects. In the kitchen, non line of sight furniture was moved. In the bedroom, kitchen and bathroom, during the collection of data set B, another person was present in the monitoring area besides the person who falls.

5.6.4 System Calibration

FallDeFi depends on input parameters at four different stages: *denoising, stream selection after PCA, spectrogram processing* and at *threshold selection for event detection*. Denoising and PCA stream selection parameters are selected only once and these parameters are not changed across environments. Spectrogram processing and event detection parameters depend on the testing environment. However, the process of initializing them is mentioned here.

5.6.4.1 Denoising

There are four parameters that must be initialized in wavelet de-noising: *the wavelet function* and its *coefficients*, the *number of wavelet levels*, and the *thresh-*

89

olding strategy. As there are several strategies to select these four parameters, they were studied heuristically to select the best strategy for CSI signals. The amount of noise in CSI is the main characteristic that changes from one environment to another. Once the strategies are selected, DWT adjusts its thresholds to lower wavelet levels by estimating the noise in the highest wavelet. Therefore, this is a one time selection and initialization for all experiment environments. The criterion for tuning those parameters are: high correlation of the synthesized signal to the original signal after de-noising, preservation of high frequency components, and computational complexity.

The synthesized signals were obtained by considering different parameter combinations and they were compared to the raw signals using the Pearson correlation coefficient. The parameters used in the synthesized signals that yielded a high correlation compared to the raw signals were selected and they were 'symlet' *wavelet function* with three *coefficients* and 10 *wavelet levels*. The wavelet function 'symlet' is chosen from other commonly used wavelet functions such as 'Haar', 'Daubechies' and 'Morlet', the considered wavelet coefficients were 2, 3, 4 and 5 and the wavelet levels were 4, 6, 8, 10 and 12. As the wavelet coefficient increases in number, the time resolution and computation time increases while the wavelet level increases, frequency resolution and computation time increase. The *thresholding strategy* used is: fixed-form threshold with threshold rescaling for individual wavelet levels and soft thresholding [129]. Fixed form thresholding is selected by comparing with other commonly used strategies like 'rigsure', 'heursure' and 'minimaxi'.

5.6.4.2   Stream Selection

PCA stream selection requires only one parameter (captured variance) to be initialized and it is fixed for all environments. The main goal of captured variance initialization is to avoid selecting noisy PCs for the spectrogram. However, depending on the noise levels in each environment the number of selected PCs will vary dynamically.
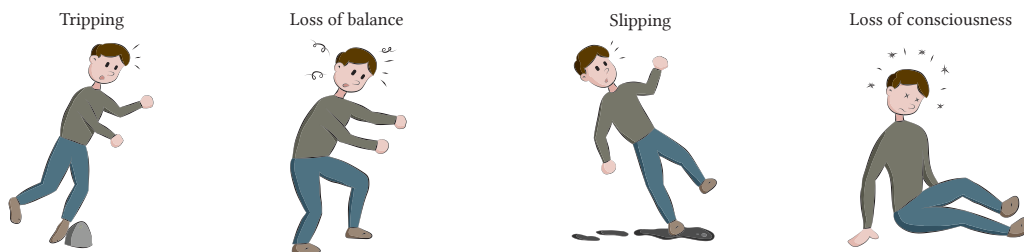
Figure 5.11: Falling Types.

Table 5.3: Performed activities in the experiments.

	Main categories	Sub variations
Fall Types	trip, slip lose consciousness lose balance	forward, backward, lateral, on-position
Other Activities	walk, jump, no event bend and pickup	
	random event	open, close door, raise hands
	sit down, stand up	floor, chair

5.6.4.3 Spectrogram Processing and Event Detection

The threshold $\hat{N}_{th}$ used for spectrogram segmentation of each ADL is estimated by considering several spectrograms corresponding to different activities for each environment. During the calibration phase $\hat{N}_{th}$ for each spectrogram is obtained from each environment and a median segmentation threshold is computed for that environment. During the testing phase, this threshold segments all the spectrograms to isolate the main activity's time-frequency components. Since the power burst curve threshold (PBC_{th}) is dependent on this threshold, event detection step is also impacted by this value.

5.6.5 Activity Types

Table 5.3 specifies the types of activities that were considered in the experiments. Some of the activities in the main category have other variations, i.e, trip, slip,

91

lose consciousness and lose balance can occur either as a forward, backward, lateral or as an on-position motion. These variations were mainly considered to avoid over-fitting to a particular falling motion. As each person may imitate the falls uniquely, to avoid over-fitting to a particular person, three volunteers aged between 27 to 30 years with different physiques performed the activities. In all experimentation environments, two combinations of those volunteers were evaluated: only one person is in the area of interest, two persons are in the area (when one person performs falling activities, the other is either involved in regular activities or is motionless). More than two volunteers were not used in the experiments due to the application scenario, elderly falls go unnoticed when the inhabitants live independently. Hence, the system should be properly tuned to the worst case scenario where a single or couple of occupants are in the area of interest.

5.6.6 Evaluation Metrics

The following metrics that are commonly used in machine learning tasks are used here to evaluate the performance:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{5.17}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{5.18}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{5.19}$$

$$\text{F1 score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{5.20}$$

where TP = true positives, FP = false positives, TN = true negatives and FN = false negatives. Note that true positives are manually annotated falls that are correctly detected, false positives are non-falls that are detected as falls, true negatives are non-falls that are detected as non-falls and false negatives are falls that are detected as non-falls. Out of these metrics, recall and precision are the most important measures for the system. Recall provides the fraction of falls that were correctly identified among actual falls by the system. Precision identifies the fraction of relevant falls among all the detected falls, therefore it is a

measure of false alarms. F1 score is the harmonic mean of these two metrics and accuracy identifies the fraction of correctly identified falls and non-falls out of all the detected events. For a good fall detection scheme all four metrics should be high, ideally close to 100 %.

5.7 Performance Evaluation

In this section the performance of FallDeFi is evaluated in different conditions, i.e., rooms, link types, persons and activities as mentioned in Section 5.6. First, the performance of individual signal processing components of the system is evaluated and their optimum operating points are identified. Then the performance of the extracted features in classifying falls, i.e, OFs and SFs are analysed as defined in Section 5.5. Unless otherwise mentioned, for the rest of the evaluations only the SFs (event duration, spectral entropy (1 to 10 Hz & 10 to 30 Hz) and fractal dimension) are used for classification, as detailed in Section 5.5.2. Finally the performance of the system is compared with two state of the art schemes (CARM [26] and RTFall [100]).

A similar pattern of high variance in the classification accuracy in all experiments was noticed as a consequence of two reasons: *a)* low sample size of the data sets and *b)* class imbalance between falls and other fall like activities. As a solution, synthetic samples were generated from the data using the Synthetic Minority Over Sampling Technique (SMOTE) [130].

5.7.1 Optimization and Validation of Different Components of FallDeFi

5.7.1.1 Validation of Noise Filtering

A wavelet based de-noising technique was introduced in Section 5.3.1. Recall that this approach enabled the first principal component (together with other principal components) in constructing the spectrogram even for noisy links. In this section the overall classification performance is compared if the first principal component is used with and without noise filtering. Figure 5.12 compares the average results
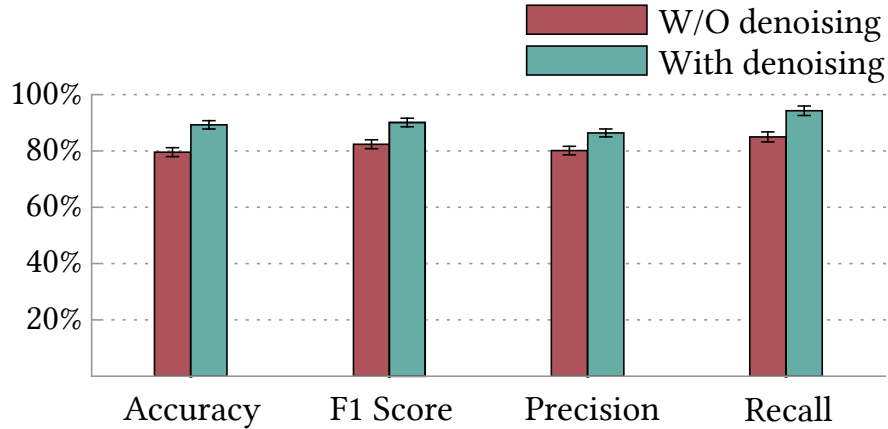
Figure 5.12: Performance with and without de-noising as mentioned in Section 5.3.1.

for all the noisy and noiseless data sets and illustrates that noise filtering improves the performance metrics by 5% to 10%.

5.7.1.2   Selection of the Optimum Number of Principal Components

As explained in Section 5.3.2, the optimum number of principal components to be used in the spectrogram is heuristic and this number is determined by the captured variance metric. Here the classification results are compared with the captured variance of CSI which ranges from 85% to 99% corresponding to the required maximum number of principal components. The 100% value for captured variance is not plotted as it does not reduce the original number of CSI streams and requires a significant amount of computational time as all the principal components have to be used in the spectrogram. As depicted in Figure 5.4a the classification performance (y axis in the figure) increases up to 95% of captured variance (x axis) and starts to reduce from thereon. This shows that choosing principal components corresponding to captured variance below 95% may not provide enough information. Moreover, using a higher number includes noisy principal components which again reduces the performance. Therefore, in FallDeFi, 95% of captured variance is selected as the optimum value for principal component selection. This amount corresponds to 2 to 3 principal components on average.

(a) Performance of the cumulative variance of selected PCs as a percentage of cumulative variance of all the subcarriers.

(b) Performance comparison at different transmission rates.

5.7.1.3 Performance of FallDeFi under Different Transmission Rates

Here the range of transmission rates for which FallDeFi performs well is evaluated. The testing is done for rates ranging from 200 pkts/s to 1000 pkts/s with increments of 200 pkts/s. Different rates were achieved by down-sampling the transmission rate of 1000 pkt/s used in all the experiments so that parameters such as activities, positions and environments are fixed and the only variable is the sampling rate. The performance, as shown in Figure 5.4b, is constant up to a transmission rate of 600 pkts/s and starts deteriorating from 400 pkts/s onward. This behavior is due to the fact that when the transmission rate is reduced, the bandwidth of the measured activity is also reduced. Thus, low frequency activities are perceived to be similar to high frequency activities resulting in lower classification accuracies. These results indicate that the performance of the system is consistent up to transmission rates of 600 pkts/s. Nonetheless, all the analysis in this paper was conducted using transmission rates of 1000 pkts/s.

5.7.2 Evaluations of the Original and Selected Features

As explained in Section 5.5 a set of STFT-based features were extracted which were termed as Original Features (OF). Then, Section 5.5.2 used sequential forward selection to select a feature set from the OFs that is robust to environment changes and were termed as Selected Features (SFs). In this section the evaluation of OFs and SFs across all the defined performance metrics is presented and the classification performance of the system is analyzed in the two groups A and B listed in Table 5.2 in three different ways:

(i) First the classification performance is assessed in training and testing on the same data set (*A,A& B,B*). For this 8 SVM models were trained for Corridor, Bedroom, Kitchen and Bathroom in groups A and B and tested on the same data set in each group by dividing them into 70% training and 30% testing.

(ii) Then the classification performance was analyzed for training on group A and testing on group B (denoted as *A,B*). Four SVM models are trained for the data (corridor, bedroom, kitchen, bathroom) in group A and tested on

the corresponding data in group B. Here, whether the models are resilient to changes that occur over time in the same environment is tested since the data in group B were collected several days after group A's data.

(iii) Finally, an SVM model is trained for the Lab data set in group A and test on each data set in groups A and B of the other four environments (denoted as *Lab, A+B*). Here, it is tested whether a model can be trained in a completely different environment and still be applied to other environments to eliminate the need for training each time a new system is deployed.

5.7.2.1 Performance of Original Features

Here the performance of 8 models on the eight testing data sets consisting of OFs are evaluated. The average results are illustrated in Table 5.5a as *OF: A,A& B,B*. The figure shows that for OFs, all the performance metrics have values above 93 %, especially, with recall reaching as high as 97.5 %. This translates to a highly reliable fall detection system when the system is trained to the prevalent environment conditions.

Then the performance of OFs for A,B and Lab,A+B as mentioned above in Section 5.7.2 are tested. The results of these two scenarios are indicated in Figure 5.5a as *OF: A,B* and *OF: Lab,A+B* respectively. As evident from the figure, all the performance indication metrics degrade considerably to a range between 55 % to 80 % for both the scenarios. This indicates that for OFs, the system performs poorly in detecting falls after environment changes occur or if the model is trained in a different environment.

5.7.2.2 Performance of Selected Features

To evaluate the performance of SFs, the same procedure as mentioned above in Section 5.7.2.1 is followed. Figure 5.5b shows the average classification results for SFs for the three scenarios, *SF: A,A&B,B*, *SF: A,B*, and (*SF: Lab,A+B*). Average results indicate that *SF: A,A&B,B* performs only 3 % worse than *OF: A,A&B,B*, yet, all metrics exceed 90 %. This is a clear indication that when trained to the prevalent conditions, the difference in performance of SFs and OFs is small. However, for the latter two scenarios, *SF: A,B* and *SF: Lab,A+B*, an

improvement in accuracy, f1 score and recall for SFs is noticed. This indicates that the SFs have especially improved the ability to detect falls of the system albeit a decrease (12 %) in precision for *SF: A,B* from *OF: A,B*. Yet, all the average performance metrics are closer to 80 % for SFs. More importantly, this illustrates that the SFs have not only good performance when trained to the existing environment, they are resilient in fall detection when environment changes occur or when the training is done in a completely different environment.

The reason for this behavior can be attributed to the type of features that were selected, i.e., event duration, spectral entropy and fractal dimension. Typically, an event duration of 2 s to 3 s is a unique characteristic of falls irrespective of the environment. Spectral entropy and fractal dimension are normalized features, the former quantifies the variations in the amplitudes in a specific frequency range while the latter quantifies the roughness of the extreme frequency curve.

5.7.2.3   Robustness of Selected Features on each Environment

Average performance results indicate the general tendencies of the system, however, they do not yield insights into the impact on each environment. Here this view is provided using only the selected features because they have the best performance in all the conditions. In Figures 5.6a, 5.6b, 5.6c and 5.6d the accuracy, f1 score, precision and recall is compared of corridor, bedroom, kitchen, lab and bathroom. First the results are shown for training and testing on the same data sets of A (*SF: A,A*) and B (*SF: B,B*), then the results of training on data set A and testing on the data set B (*SF: A,B*) are compared. Finally the results for training on Lab data and testing on both A and B data sets (*SF: Lab,A+B*) are compared.

The results for individual environments confirm the average results in Section 5.7.2.2. Even though training and testing on the same data set performs slightly better than the other two scenarios as expected, all the metrics perform above 70 %, especially, the accuracy, f1 score, and precision are either above or closer to 80 %. For recall, the lowest performance is ≈70 % in corridor and bathroom data sets for *SF: A,B*. Generally, for corridor bedroom and bathroom, other metrics also perform slightly lower than for kitchen. The reason is that for the former three environments, the link is not as strong as in the kitchen (corridor:
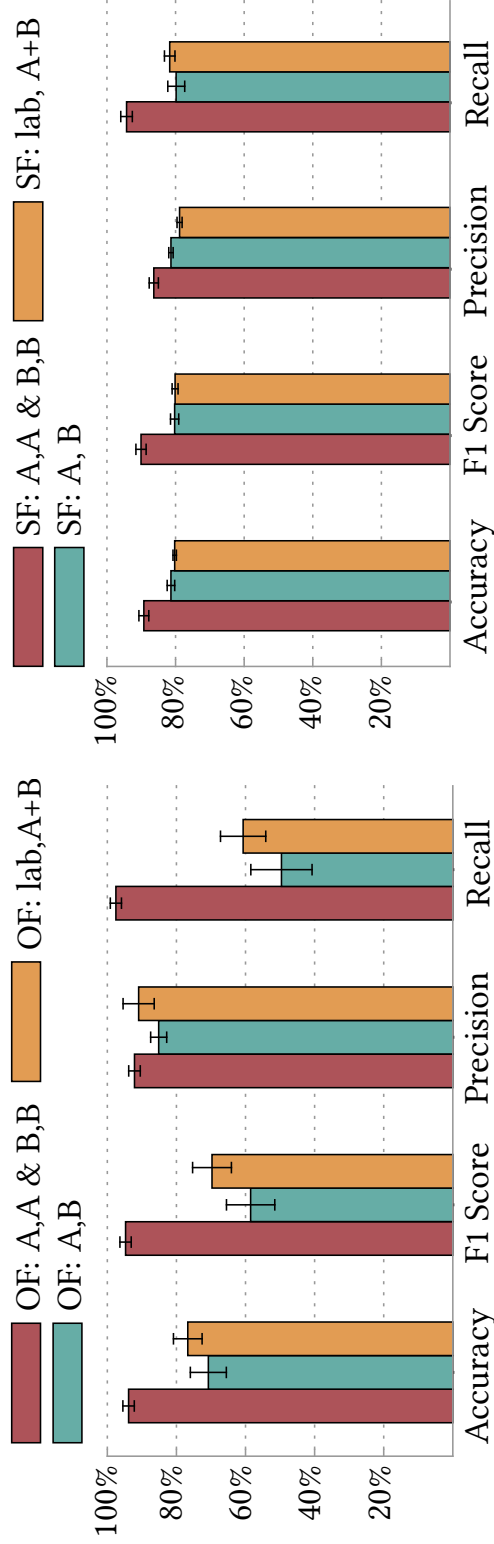
(a) *OF*–Original features, *A,A&B,B* training on A and B and testing on A and B. *A,B*–training on A and testing on B. *Lab,A+B*–Training on Lab data set and testing on A and B.

(b) *SF*: *A,A&B,B*–Selected features, train and test on same data set. *SF*: *A,B*–Selected features, train on A and test on B. *SF*: *Lab,A+B*–Selected features, train on Lab data set and test on A and B.

Table 5.5: Accuracies of original (a) and selected (b) features.
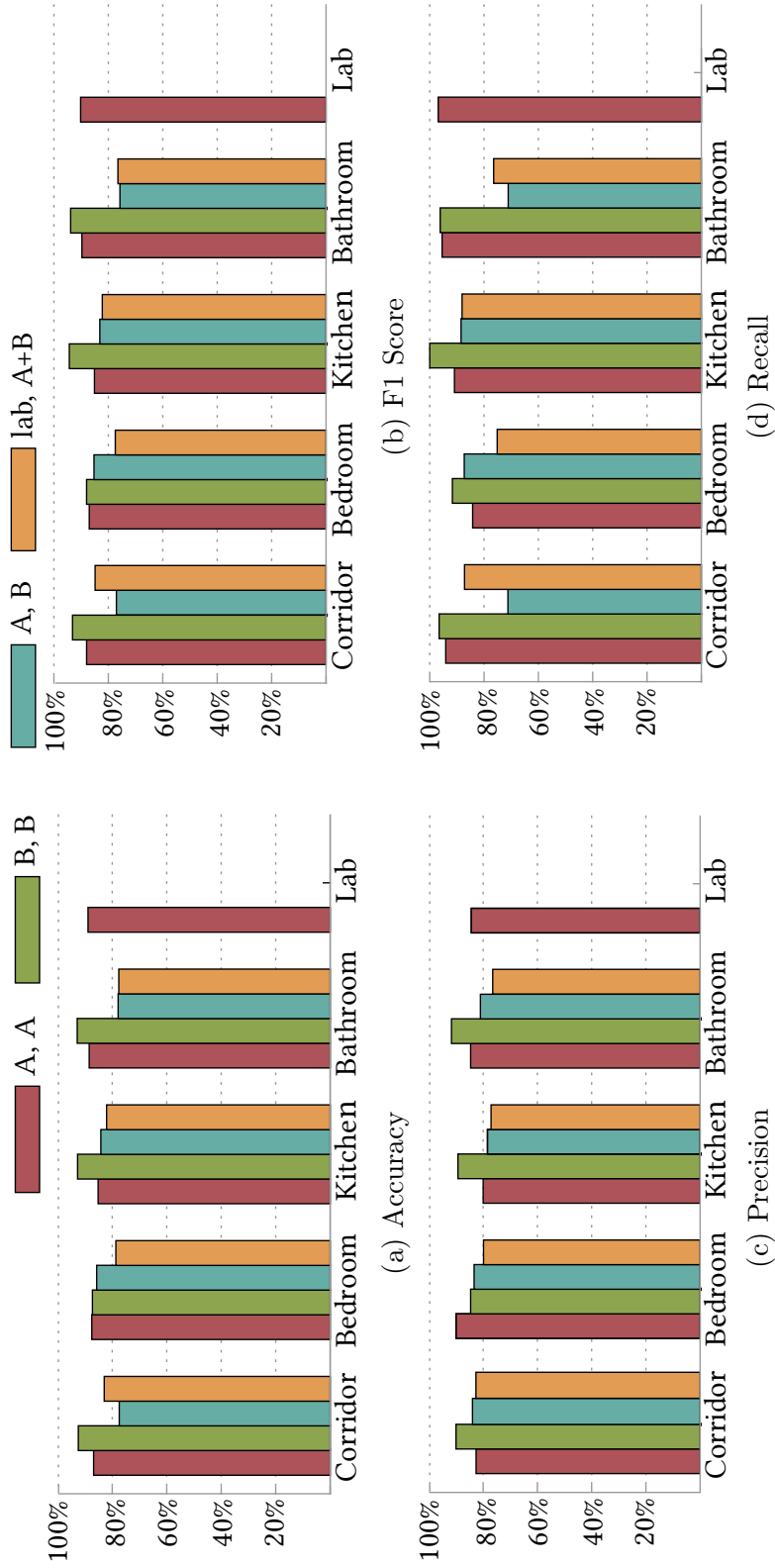
Table 5.6: Comparison of FallDeFi performance across different environments. For these comparisons only SFs are used. The notation used in all the figures are as follows. *A,A*−Training on A and testing on A. *B,B*−Training on B and testing on B. *A,B*−Training on A and testing on B. *Lab,A+B*−Training on Lab data set and testing on A and B.
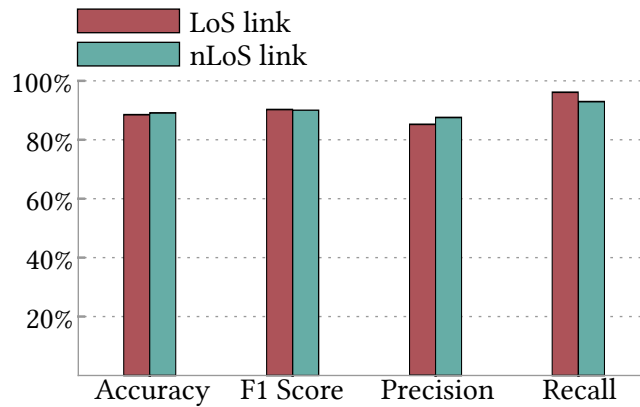
Figure 5.13: Performance comparison for falls and other activities in LoS and nLoS links.

9m LoS, bedroom: 5m through wall and bathroom: 5m through wall vs kitchen: 4m LoS). The link in the bathroom penetrates two walls while the link in the bedroom penetrates one wall. For falling positions that are far from the receiver or the link, especially when the link is weak as in above three environments, high frequencies/energies are not effectively captured from an activity. Additionally, when a fall or a similar activity e.g. sit down is closer to the link or the receiver, the amplitudes in high frequencies increase, causing an adverse effect on distinguishing falls from other activities.

5.7.3 Impact of the Link type and the Person who does the Activities

Next, the results of all the environments were divided according to the link type, LoS and non-LoS links. The results are shown in Figure 5.13. According to the results, the difference between the system's performance for LoS and non-LoS links is quite low. Recall of the LoS links is slightly higher than in non LoS links, however, precision is slightly lower than the latter. This indicates that in LoS links, detection of falls is slightly better than in non LoS links and at the same time, they produce more false positives. This is expected behavior as LoS links are more stronger, they can detect more falls. The accuracy is slightly lower in LoS links because there are more ADLs than falls in the testing set, generating
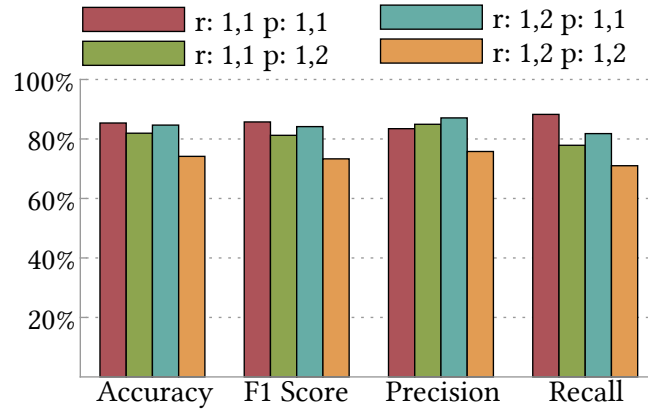
Figure 5.14: Performance comparison for training and testing in the same room for the same person (r:1,1 p:1,1), two different rooms but same person (r:1,2 p:1,1), same room but different persons (r:1,1 p:1,2), and two different rooms each having a different person (r:1,2 p:1,2).

more false positives.

Recall that three persons (p1, p2 & p3) were used to train the system. Here, how much of an impact each person's body type and unique falling characteristics would have on the overall performance of the system is analyzed. First, FallDeFi was trained and tested on for falls and other activities data collected from

a) same room, same person (r:1,1 p:1,1)

b) trained on one room and tested on another, same person (r:1,2 p:1,1),

c) same room, trained on one person and tested on another (r:1,1 p:1,2),

d) trained for a person in one room and tested on another person in a different room (r:1,2 p:1,2).

The results are illustrated in Figure 5.14 and as expected, they indicate that training and testing for the same person yields higher performance than training on one person and testing on another person. However, the accuracy of r:1,2 p: 1,1 is closer to r: 1,1 p: 1,1 indicating the effect of environment is not adversely affecting the system. The lowest performance is yielded for r:1,2 p1:2 where both the person and the environment change from the training phase to the testing phase (the accuracy is ≈10% lower than r:1,1 p:1,1). Therefore, this indicates

that more than the environment, the person the system is trained on has an adverse impact on successful detection using the current set of features. This is expected, because different physiques and different behavioral patterns of people create time-frequency signatures unique to each person.

5.7.4 Comparison of FallDeFi with CARM and RTFall

The performance of the approach using SFs is compared with two other state of the art CSI-based schemes CARM [26], an activity recognition scheme and RTFall [100], a fall detection system. CARM was selected because FallDeFi is inspired by the signal processing techniques used in CARM and it is the most comparable activity recognition scheme to FallDeFi. Even though there are two other CSI-based fall detection schemes in the literature, RT-Fall was selected specifically because it builds upon the features used by those approaches.

5.7.4.1 Implementation of CARM

The features of CARM are extracted from 12 frequency levels and the extracted features are

a) The energy in each level representing the movement intensity in each speed range.

b) Differences in the energy of each level between consecutive 200 ms intervals representing the rate of change of the energy for each activity.

c) Estimated torso and leg speeds using the percentile method [79].

Altogether 27 features were collected. The feature space of CARM is modelled using a Gaussian mixture model and the state space using a HMM. For fair comparison with FallDeFi, the event detection method of FallDeFi is used to detect events. Two HMMs are trained for falls and non-falls for binary classification. Following notation is adapted to indicate the complete parameter set of the HMM model: *a)* hidden (unobserved) states $\boldsymbol{S} = \{before fall, fall, after fall\}$, correspond to three different behaviors of a human activity ; *b)* initial state probabilities $\boldsymbol{\pi}$; *c)* observations $\boldsymbol{O} = \{o_1 \dots o_{27}\}$, correspond to the features used to

characterize the interference *d)* state transition probability matrix $\boldsymbol{A}$, models the evolution of the activity states as transitions among the set of unobserved states; *e)* observation probability matrix $\boldsymbol{B}$.

The defining parameters of the Gaussian Mixture Model (GMM) are the number of components ($\mathcal{M}$) and three matrices: mixture component weights ($\boldsymbol{W}$), component means ($\boldsymbol{\mu}$) and covariances ($\boldsymbol{\Sigma}$). The former is a stochastic matrix which determines the weight at which each Gaussian component should model data, and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ define the mean and the covariance of each component. In the current approach 7 GMM models were selected through cross validation. The choice of the number of components ($\mathcal{M}$) affects the estimation accuracy. Moreover, each component ($\mathcal{M}$) has ($\mathcal{Q}$) dimensions given by the number of features used to characterize the distributions.

The mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of the Gaussian mixture, state probabilities and transition probabilities of $\boldsymbol{A}$ and $\boldsymbol{B}$ of the HMMs are initialized as mentioned in [26]. To guess the initial values of $\boldsymbol{A}$ and $\boldsymbol{B}$, CARM divides the sequence of feature vectors from each activity in the training set equally among the states and then calculates the initial values for the mean vector and covariance matrix of each state using the feature vectors assigned to that state. To estimate mean vector and covariance matrix corresponding to each state and the transition probabilities for the HMM, CARM uses the Baum-Welch algorithm.

In addition, the training data used for the HMM is labeled as before fall, fall and after fall through visual inspection. In this approach, two HMM models are used, one for falls and one for non-falls. Once the parameters are computed and set, the model is trained. The matrices (i.e. $\boldsymbol{W}$, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$), of the GMM model were estimated using the EM algorithm. A diagonal covariance matrix $\boldsymbol{\Sigma}$, the most used in the literature was adopted, requiring less samples for training, and approximating full covariance using a linear combination of diagonal covariances.

5.7.4.2 Implementation of RTFall

RT-Fall is implemented exactly as described in the literature using CSI phase difference from two antennas to extract features. The extracted features of RT-Fall from CSI phase difference of a subcarrier are as follows.

a) the normalized standard deviation,

b) the median absolute deviation,

c) the offset of signal strength,

d) Interquartile Range (IR),

e) signal entropy,

f) the rate of change of the signal - computed as the first order derivative which is an indicator of how fast the signal is changed due to an activity

g) the TimeLag - defines the time delay of the state transition between the band-pass filtered and the raw phase difference

h) the Power Decline Ratio (PDR) -

$$PDR = \frac{\sum_{t=t_1-1}^{t_1}\sum_{f=f_1}^{f_2} e_{t,f}\omega_f - \sum_{t=t_1}^{t_1+1}\sum_{f=f_1}^{f_2} e_{t,f}\omega_f}{\sum_{t=t_1-1}^{t_1}\sum_{f=f_1}^{f_2} e_{t,f}\omega_f} \tag{5.21}$$

where t_1 is the time at finishing point. f_1 and f_2 refer to the frequency range from [0,50Hz]. $e_{t,f}$ is the power of frequency f at time t. ω_f is the weight at frequency f.

The CSI packet transmission rate is set to 100 pkts/s which is achieved by down-sampling the 1000 pkt/s rate. The band-pass filter had a range of 5 Hz to 10 Hz. As the feature, time lag relies on RTFall's event detection method, it was also implemented. SVM was used for binary classification and the two thresholds for event detection were determined from 5-fold cross validation.

5.7.4.3 Results Comparison

In Table 5.7, the average performance of FallDeFi, CARM and RT-Fall when the systems are trained for the same data set and tested on the same *(A,A&B,B)* is first compared. Then a separate model for each environment in data set A is

trained and tested for each model for the corresponding environment in data set B (*A,B*). In this comparison, whether the three schemes are resilient to changes that occur over time in the same environment is tested since the data in group B were collected several days after group A's data. Then a model is trained for the Lab data set and tested for both the data sets A and B in corridor, bedroom, kitchen and bathroom (*Lab,A+B*). Here, it is compared among the three schemes whether a model can be trained in a completely different environment and still be applied to other environments to eliminate the need for training each time a new system is deployed.

The performances are measured using accuracy, F1 score, precision, recall/true positive rate (TPR) and false positive rate (FPR). Recall/TPR characterizes the number of false alarms generated in relation to missed fall events

$$FPR = \frac{FPs}{FPs + TNs} \tag{5.22}$$

characterizes the false alarms at zero-fault level. These metrics are average values for a total of 326 falls and 744 non-falls.

In Table 5.7 these values are compared with the state of the art. For training and testing in the same environment, results show that the accuracy of FallDeFi is 11% better than CARM and 8% better than RTFall, the TPR is 14% better than CARM and 9% better than RTFall and the FPR is 10% less than CARM and 15% less than RTFall. FallDeFi is also more robust across different environments than CARM and RTFall as indicated by the *A,B* and *Lab,A+B* results. A notable trend among the three systems is that they all perform better when tested in the environments they are trained in. RTFall suffers the most, with 20% and 30% decrease in accuracy in A,B and Lab, A+B respectively. FallDeFi fairs the best among the three schemes, with just 8% decrease in accuracy in both A,B and Lab, A+B. The reasons for this performance enhancement compared to other two schemes can be attributed to the denoising technique and the resultant activation of the 1st PC (1st PC contains the majority of information in the subcarriers) and the rigorous feature selection procedure that was followed to extract environment resilient features.

However, this performance comes at the price of computational time. The data

in Table 5.8 illustrates this. FallDeFi takes twice the computational time (0.6s) than RTFall (0.34s), especially due to feature extraction from the spectrogram. These computation times are for 1 s of measured data. Therefore, the important fact is that even with the added computation efforts, FallDeFi is still able to produce results before the next second of data measurements, offering real-time execution. Note that these performances are achieved in a laptop having an intel Core i7 processor at 2.7 GHz and 16 GB of RAM.

5.8 Discussion and Conclusions

In this chapter a fall detection system comprising of commodity WiFi devices is presented. A novel methodology for time-frequency analysis of CSI for WiFi fall detection including improved CSI preprocessing techniques is provided. Using these techniques a high performance was obtained under transmission rates as low as 600 pkts/s. A pre-screener was developed using the spectrogram and was used for feature extraction. Through effective feature extraction in time-frequency domains, falls can be detected and classified with low dependence on the training environment. Above 93 % average accuracy can be achieved for a pre-trained system while the average accuracy is closer to 80 % with changes in the environment or when the system is trained in a different environment. An overall TPR of 80 % and an FPR of 20 % is achieved when the system is trained in a different environment. This translates to 2 in 10 falls are undiscovered and 2 in 10 non-falls are detected as falls. Additionally, it can be assumed that in situations where falls occur at locations with low WiFi coverage or in weak links, the TPRs further decrease (Fig. 5.6d). As the majority of events that occur during a day in real life are non-falls, FPRs in excess of 20 % are still not adequate which may result in annoying the users involved in daily activities. These two shortcomings can be addressed using signal processing steps and performing further experiments to see the effect of position of the fall. Currently, 5 access points are chosen for a $60 \, m^2$ area (Fig. 6.9a) and placed them in positions that were perceived to achieve maximum coverage. Therefore, an assessment of the optimum position of WiFi access points for a successful deployment in domestic environments is required. These enhancements are planned to be addressed in the next chapter.

Table 5.7: Average classification comparison of FallDeFi, CARM and RTFall for Corridor, Bedroom, Kitchen, Bathroom and Lab. Used notation: $A,A\&B,B$–Training and testing on same data of A and B; A,B–Training on A and testing on B; $Lab,A+B$–Training on Lab data and testing on A and B. The results of the two best schemes for each data set are highlighted in bold.

%	A,A & B,B			A,B			Lab, A+B		
	FallDeFi	CARM	RTFall	FallDeFi	CARM	RTFall	FallDeFi	CARM	RTFall
Accuracy	**88.90**	77.85	**80.96**	**80.90**	75.68	59.35	**80.10**	64.26	51.14
F1 Score	**90.24**	76.41	**82.45**	**82.34**	77.15	57.72	**80.15**	58.34	57.00
Precision	**86.65**	76.73	**80.01**	76.12	73.60	56.63	**78.95**	**67.90**	51.40
Recall/TPR	**94.33**	80.32	**85.20**	**90.48**	81.70	62.81	**81.56**	53.07	**67.82**
FPR	**14.92**	24.63	**24.00**	27.87	30.34	45.52	21.56	**24.55**	65.91

Table 5.8: Comparison of FallDeFi, CARM and RTFall processing times for one second of CSI.

	CPU time					
	FallDeFi		CARM		RTFall	
	(s)	(%)	(s)	(%)	(s)	(%)
Amplitude/phase extraction	0.282	45.30%	0.282	46.70%	0.290	85.00%
Denoise	0.018	3.00%	0	0.00%	0.001	0.29%
PCA	0.002	0.30%	0.002	0.33%	0	0.00%
Spectrogram	0.020	3.20%	0.020	3.30%	0	0.00%
Feature Extraction & classification	0.300	48.20%	0.300	49.67%	0.050	14.66%
Total	**0.622**		**0.604**		**0.341**	

As mentioned in Section 5.6.3, a data rate of 1000 pkt/s, which is equivalent to 800 kbps is used. As the results in Section 5.7.1.3 show, this can be further reduced to 480 kbps without a degradation in accuracy. This amount of data rate is easily achievable in existing WiFi networks, especially 802.11n devices without degrading the network performance, when mixed with other services. However, it is envisaged that this fall detection system can also be installed as an independent system from existing home wireless networks as the cost of WiFi access points and receiver cards is quite low compared to other ambient fall detection systems such as Doppler radars.

Although CSI is available in all 802.11n compatible chipsets, it can be accessed through software only in a limited number of chipset ranges with modified drivers and firmware e.g. Intel 5300 and Atheros 9k series. However, this issue is not just limited to this system. With the current expansion of research using CSI as the signal strength descriptor, chipset manufacturers are expected to expose CSI externally in a wider range of products.

As it is not possible to subject elderly people to simulate the falls, the data were collected from young volunteers (aged between 27-30 years). This implies that the system parameters require tuning for elderly that fall in real world situations. This issue is also common to all fall detection systems. Nonetheless, further analysis is required for adaptation mechanisms of the system to such scenarios.

Chapter 6

Enhancements to Fall Detection

In this chapter, FreeFall, a non-wearable fall detection scheme using CSI with commercial WiFi devices is proposed. FreeFall is an extension of FallDeFi presented in the previous chapter. In this chapter, the main challenge faced in FallDeFi is addressed which is improving the accuracy and reducing the false alarm rate to facilitate real world deployment. Additional experiments are also performed to identify the impact of external factors on detection accuracy other than the WiFi signal such as fall position, link type and room size.

Therefore the main contributions of this chapter can be summarized as:

a) Segmentation of activities to extract representative features and fall related feature extraction,

b) CSI amplitude sanitation to remove artifacts from CSI, and

c) analyzing the impact of fall position, the impact of access point position and room size on the accuracy of fall detection.

The key research challenges addressed in FreeFall and the detailed contributions of this chapter are as follows.

a) In real life, a fall is not always an isolated event. A fall can be preceded and/or followed by other human activities such as walking, standing up or sitting down. Traditional fall detection/activity recognition systems initially detect a human event through a motion detection method and then classify the activity in that event assuming a single activity occurs during the

detected event duration. This can lead to misclassification because the extracted features contain information from a mixture of events. This problem is addressed by first detecting the event, then segmenting the individual activities in that event and finally identifying a possible fall in that event period. Human velocity profiles are used to segment different activities and extract features from these segmented activities.

The next challenge is how to segment the individual activities in an event. The understanding of human motion is used to solve this. Each human activity during an event period can be characterized by a series of increases and a decreases in velocity due to acceleration and deceleration of the body. A fall fits perfectly to the above description as it is characterized by a sudden large increase in velocity due to gravitational force followed by a sudden decrease due to the body hitting the ground. For segmenting a fall from other activities, the transitions in velocity are captured using the local minima and maxima in the velocity profiles.

New features are extracted from these segmented activities and feature selection is used to enhance the resilience to environment changes.

b) CSI amplitude and phase contain the Doppler shift, which is caused by human body motion. Human activity recognition studies extract the change rate of the signal propagation path length from CSI amplitudes through signal processing. Change rate of the path length is a measure of the Doppler shift, thus, it is also related to the velocity of human movement [26]. However, CSI is still corrupted by other artifacts such as signal attenuation due to line of sight movement, unwanted frequency shifts from different body parts and other radio frequency interference. In this work these unwanted interferences are eliminated from change rate of path length using a combination of in-band noise filtering and a novel principal component selection algorithm is used to remove the artifacts in low frequencies caused by multipath propagation, line of sight attenuation and other interference terms. This helps to obtain precise human velocity profiles from the spectrogram.

c) Experiments are carried out in three different environments to test robust-

ness under different conditions and analyze the impact on fall detection from falling position, distance from the link and the impact of the size of the room. In all the experiments, commercial WiFi devices are used to extract CSI as already mentioned in previous chapters. Further experiments are run in five different environments in LoS and non LoS, through-wall, cluttered and uncluttered conditions. Empirical results from over 300 falls using 3 human subjects show that FreeFall has 94.7% average accuracy, 96.33% True Positive Rate (TPR), and 5.63% False Positive Rate (FPR) when tested in the training environment. When the testing environment is different from where the system was trained, FreeFall still achieves 92.98% average accuracy, 88.59% TPR, and 6.14% FPR. This is an improvement of 16% in accuracy, 8.6% improvement of TPR and 15.42% improvement of FPR compared to the next best fall detection scheme. Compared to an accelerometer based fall detection system executed for the same set of falls, this system has a 2.6% decrease in accuracy 4.03% decrease in TPR and 2.3% decrease in FPR.

6.1 Characterization of a Fall seen by WiFi

In this section, time-frequency behavior of a fall captured by CSI is analyzed. Fig. 6.1a illustrates a fall where a volunteer falls on to a mattress. The figure is divided into 5 zones from (a) to (e) depending on the motion pattern of the human.

Zone (a) is the phase where ADLs occur. In the given example in Fig. 6.1a, only low frequency activities ($< 20Hz$) can be observed in this zone. In zone (b), a slight dip in energy can be seen which characterizes the transition from one ADL to another through a silent period. (Same behavior can be observed in other experiments that involve ADLs other than falls where transition from one ADL to another generates a dip in energy due to low activity levels.) Zone (c) characterizes the free falling stage where there is high amount of energy in high frequencies due to acceleration of the body. The point of highest frequency characterizes the impact point on the floor and in the experiments, the highest frequency varies from $30Hz$ to $180Hz$. The wide range for the highest frequency depends on a number of factors, e.g. speed of the fall, distance of the fall from
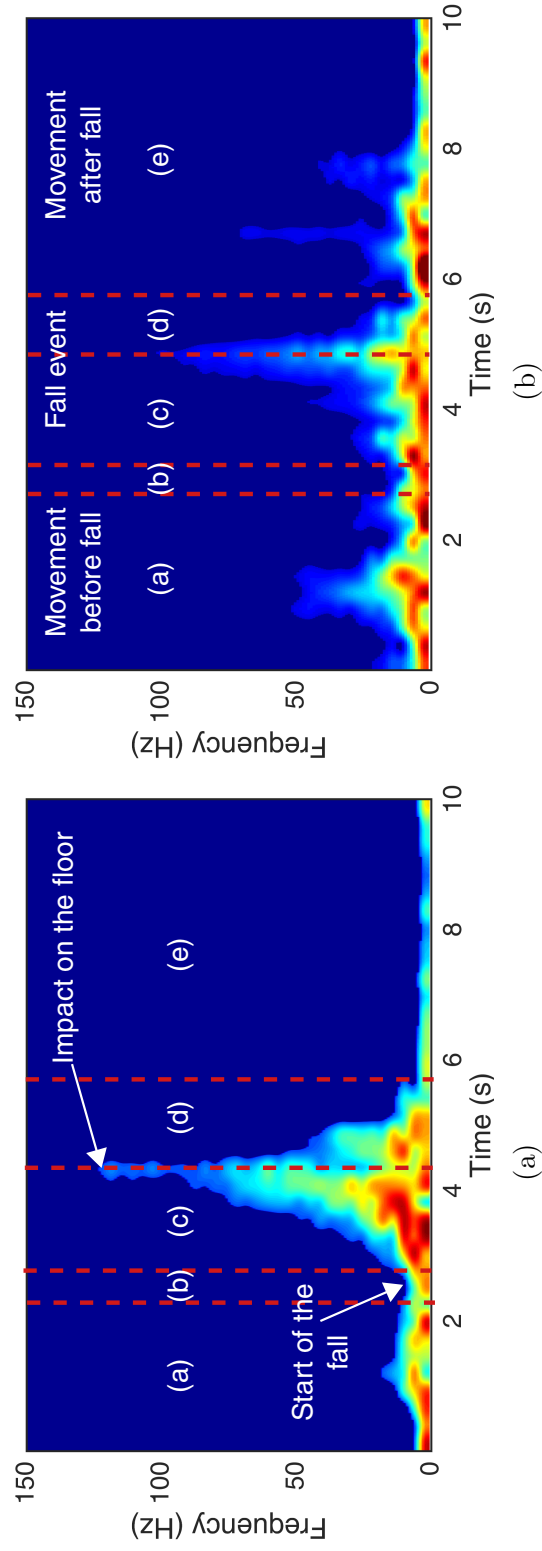
Figure 6.1: Anatomy of a fall. (a) A pure fall in a detected event. (b) A fall and other ADLs in a detected event.

the link and link type. The impact of these factors in fall detection accuracy is shown in Section 6.5.1. In zone (d) a sudden drop in energy in high frequencies can be seen due to the sudden deceleration of the body. At this point the person sinks into the mattress and there could be minor movements in the body as the person settles down on the mattress and in zone (e) the person is lying still.

Although Fig. 6.1a shows a pure fall, falls that occur during activities of daily living do not necessarily follow the same pattern. As an example, a fall can occur while walking due to slipping or tripping over an object. Additionally, the person can recover from the fall and can also stand up. Fig. 6.1b characterizes such a situation where there is movement corresponding to frequencies up to $50Hz$ before and after a fall. Existing event detection schemes [14; 26; 86; 99; 100] cannot localize the fall amid other activities, instead they detect a composite event from the event detection method consisting of the fall and other ADLs and extract features from the whole event. Section 6.2 explains how this problem is overcome.

6.2 Isolation of a potential Fall from other Activities

Isolation of one ADL from another is not a trivial task in the time-frequency domain. As a naive approach, a fall can be isolated from ADLs setting a high frequency as a threshold to detect energies in high frequency regions, but other ADLs having similar patterns to falls such as sitting down quickly can cause false alarms. Additionally, in movements such as where a person falls from a chair may not produce high frequency components at all. As the person is falling from a lower height when falling from a chair than standing, this can translate to a low velocity event [131]. To address this issue, the knowledge on relative highs and lows of energy levels during the transition from one activity to another in the time frequency domain is used here. Note that in both Figures 6.1a and 6.1b the transition from ADL to fall can be identified in zone (b) where there is a small consolidation period characterized by low energies. This is due to deceleration of the body from the previous ADL, momentary silence followed by an acceleration
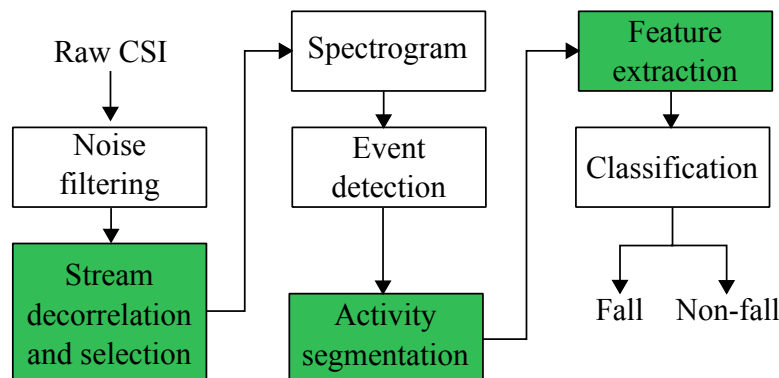
Figure 6.2: *FreeFall* system architecture. The novel contributions of FreeFall are highlighed in green.

(fall). This is captured in the CSI by a variation of high-low-high in the signal propagation path length (in time-frequency domain this is reflected as a transition from high-low-high frequency components).

The variation of energy in the frequency components at a particular time can be captured by a single metric, the velocity curve [26] which provides a measure of the instantaneous frequency at a particular time and this velocity is proportional to Doppler frequency (without the movement direction) at that moment. This curve is used to segment the ADLs during transition periods.

6.3   FreeFall System Architecture

This section first explains how the raw CSI streams are processed to obtain an informative spectrogram, followed by how the velocity profile from the spectrogram is extracted for segmenting the ADLs when an event is detected. This helps to extract features representing the real fall and finally perform classification to distinguish falls from ADLs. The complete system architecture is shown in Figure 6.2. Parts that are marked in green highlight where the specific contributions of *FreeFall* are located.

Algorithm 3: *FreeFall*, fall detection algorithm

    input : CSI_{raw} - CSI measurements of a t_{int} second interval from
           $a \times b$ antenna pairs
    output: Fall detection

29 **Initialization :** $\hat{N}_{th}, PBC_{th}$;
30 **foreach** *antenna pair* **do**
31   **foreach** *subcarrier* **do**
32 $CSI_{interp} \leftarrow$ LinearInterpolate (CSI_{raw});
33 $CSI_{wden} \leftarrow$ WaveletDenoise (CSI_{interp});
34   **end**
35   PCs $\leftarrow$ PCA (CSI_{wden});
36   selected PCs $\leftarrow$ SelectOptimumPCs (PCs);
37   **foreach** *selected PC* **do**
38 S_{raw}(n,k) $\leftarrow$ ComputeSpectrogram (selected PC)
39   **end**
40 **end**
41 $S(n,k) \leftarrow$ AverageAllSpectrograms $(S_{raw}$(n,k)$)$;
42 $v(k) \leftarrow$ ConstructVelocityCurve $(S(n,k))$;
43 $PBC(n) \leftarrow$ ConstructPowerBurstCurve (S(n,k)) ;
44 **if** $PBC > PBC_{th}$ **then**   // Fall-like event detection
45   peaks, troughs$\leftarrow$DetectPeaksTroughs(v(k));
46   ADL $\leftarrow$ SegmentADLs $(v(k)$,troughs, $S(n,k))$;
47   **foreach** *ADL* **do**
48   feature vector $\leftarrow$ ExtractFeatures $(PBC(n), S(n,k))$;
49 $fall, nonfall \leftarrow$ RandomForest (trained model, feature
        vector)   // Classify falls and non-falls
50   **if** *fall* **then**
51   ADL $\leftarrow$Wait60sForADL();
52   **if** *ADL* **then**
53   nonfall$\leftarrow$fall;
54   **end**
55   **end**
56   **end**
57 **end**

6.3.1   CSI Preprocessing

6.3.1.1   Signal strength interpolation

Due to weak links from non line of sight connections and wall penetration in indoor environments, packet losses of $\approx 1\%$ can occur in the worst case scenario but in the majority of cases this is $\approx 0.1\%$. As mentioned in Section 5.3, to obtain a fixed sample rate for all the data traces, linear interpolation is performed on CSI amplitudes in all subcarriers. The packet transmission rate is set at 1000 Hz which is well above the rate required for a fall detection (the Nyquist sampling frequency of a fall is $\leq 360Hz$ at 5.2 GHz and even lower at 2.4 GHz), and then interpolate signal strengths on individual subcarriers as a solution for the loss of information due to packet loss.

6.3.1.2   Subcarrier de-noising

Since there is noise in CSI subcarriers in all the frequency bands where useful information of a fall exists, *FreeFall* uses Discrete Wavelet Transform. Through parameter optimization, DWT can be used to eliminate in-band noise, preserve high frequency amplitudes and reduce distortions to the signal as already mentioned in Section 5.3.1. There are four parameters that must be initialized to optimize the wavelet de-noising and this is explained in Section 6.4 under Parameter Initialization.

6.3.1.3   Subcarrier Decorrelation and Stream Selection

An antenna pair provides 30 streams of CSI from 30 subcarriers and they resemble correlated movement when there is activity. Previous work used PCA as a noise filtering and decorrelation technique and selected principal components that represent *the rate of change in signal propagation path length* [26] due to body movements. An addition to this method is provided here by suppressing the low frequency energies embedded in principal components using a principal component selection strategy such that high frequency amplitudes are pronounced.

CSI subcarrier amplitudes inherently have a high amount of energies in low frequencies ($< 20Hz$) and their variations contribute to the majority of the vari-

117
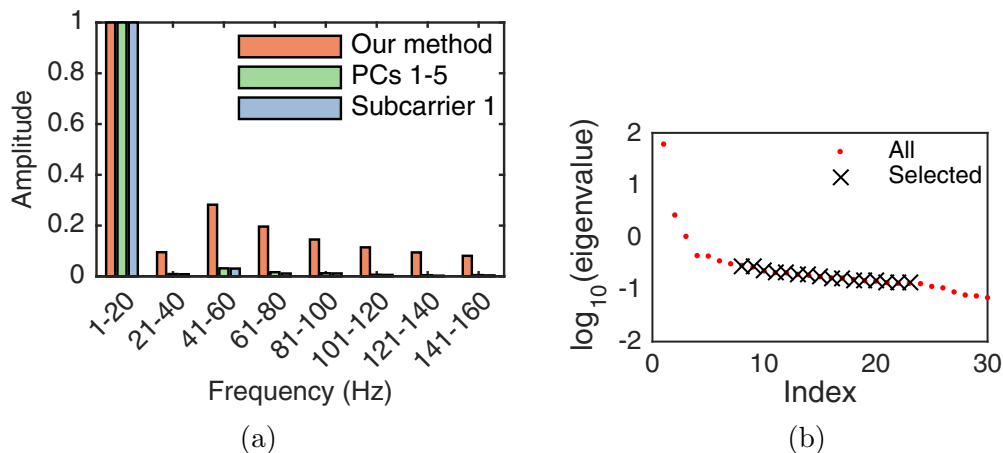
Figure 6.3: a) Average amplitude of a fall in different frequency bins normalized to the amplitudes of $1 - 20Hz$, b) Eigenvalues of the principal components of a fall

ance in the signal. Low frequency variations are predominantly due to multi-path effects, amplitude attenuation, irrelevant cross-terms, unwanted frequency shift terms and time-variant power terms due to taking the conjugate complex of $H(t, f_i)$ to eliminate the phase [102]. The high frequency elements in CSI are caused by human motion and these elements have low amplitudes because the corresponding signals are weak due to scattering off humans.

Once PCA is applied on subcarrier streams to reduce the dimensions in order to capture change rate of the path length due to human movement, in the first few principal components, it can be observed that the energies of low frequency elements $(1-20Hz)$ are higher compared to frequency elements beyond $1-20Hz$. As an example, Figure 6.3a illustrates the average relative amplitudes of a fall at 8 different frequency levels $(1 - 20Hz, 21 - 40Hz, 41 - 60Hz, 61 - 80Hz, 81 - 100Hz, 101 - 120Hz, 121 - 140Hz, 141 - 160Hz)$ having a bin size of $20Hz$ for three different spectrograms:

(i) spectrogram constructed by averaging the spectrograms of the **middle ranked principal components (5-25)** ,

(ii) spectrogram constructed by averaging the spectrograms of the **first five principal components (1-5)**,

(iii) spectrogram of a raw subcarrier.

Amplitudes in each frequency bin in the figure is constructed by averaging the energy in the corresponding frequency range of the spectrogram. It is evident from the figure that selecting the principal components in the middle ranks amplify the amplitudes of higher frequencies. Therefore, the high frequency elements in CSI have a higher impact on the middle ranked principal components. Figure 6.4 illustrates how middle rank principal components help in constructing a spectrogram having high amplitudes for high frequencies (6.4e) compared to the spectrogram of a raw subcarrier (6.4a) or the spectrogram from the first five principal components (6.4c).

Therefore in FreeFall, middle ranked principal components are selected to construct the final spectrogram for feature extraction. The process of selecting the middle ranked principal components is as follows. The middle rank principal components are selected based on their eigenvalues as eigenvalues represent the variance of the signal in the respective principal component. As shown in Fig. 6.3b, there are high amplitudes in the first seven eigenvalues, middle ranked ones (8-23) have relatively similar amplitudes and a decay in the lower ranked ones (24-30). Note that in Fig. 6.3b the eigenvalues are drawn in log scale, therefore, in linear scale the amplitude spread is much higher than shown in the figure. The difference in amplitudes in the first seven and the last seven eigenvalues compared to the middle ones is exploited to isolate the middle rank eigenvalues. For this purpose, an outlier detection technique is used, i.e the interquartile range [132]. All the principal components that correspond to eigenvalues less than the 1st quartile and higher than 3rd quartile are treated as outliers and they are eliminated. The remaining principal components are used for constructing the spectrogram.

6.3.1.4 Spectrogram Processing

For feature extraction, each selected principal component is transformed into the time-frequency domain using STFT and the average of the spectrograms to construct a single spectrogram is computed. Among various time-frequency feature extraction methods, STFT spectrogram features have performed the best
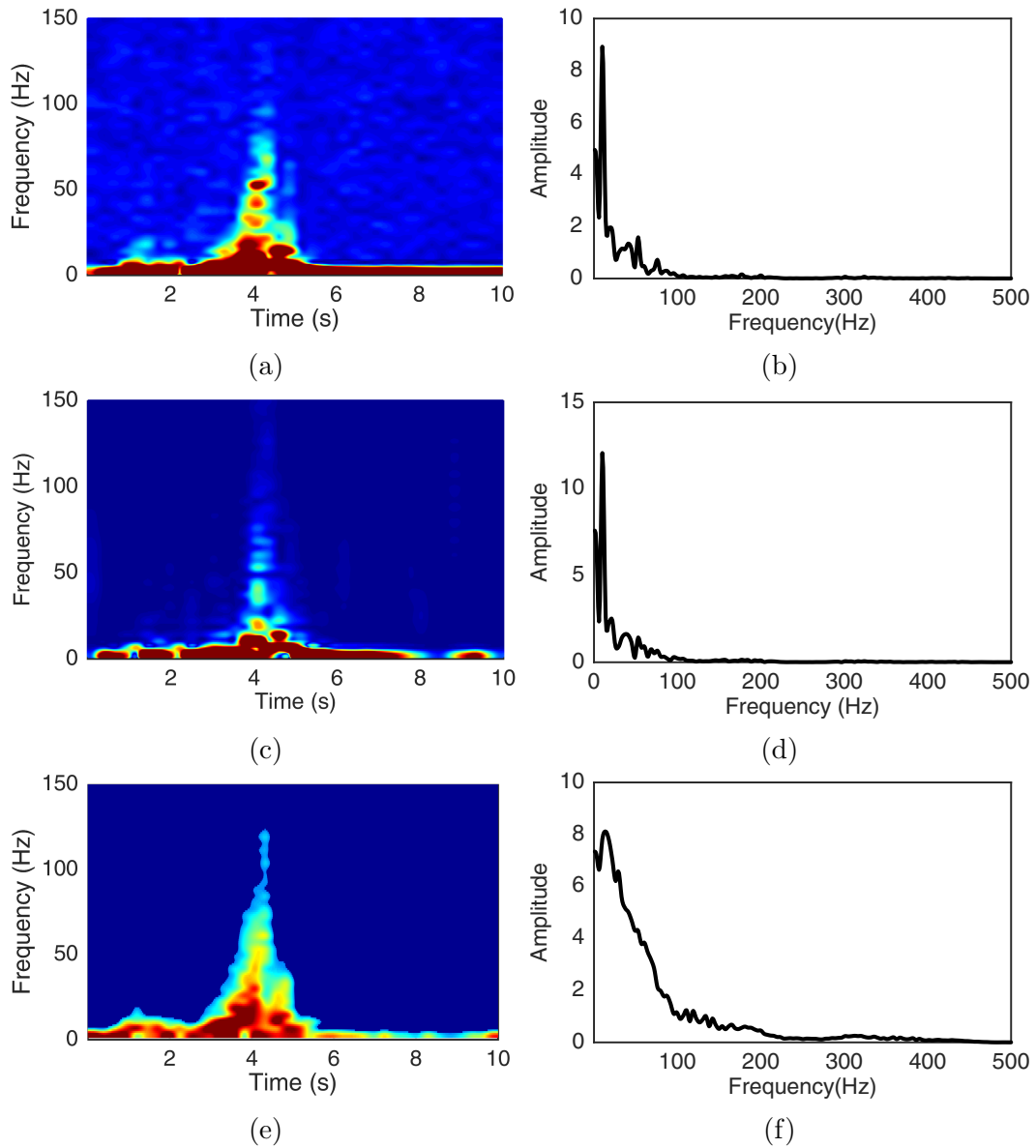
Figure 6.4: (a) Spectrogram of a single subcarrier, (b) Cross section of the spectrogram of the subcarrier at 4s, (c) Spectrogram of first 5 principal components, (d) Cross section of the spectrogram of first 5 principal components 4s, (e) Spectrogram of middle ranked principal components, (f) Cross section of the spectrogram of middle ranked principal components at 4s.

in recent radar fall detection research [77], which motivates its use here. Finally spectrogram thresholding is performed to isolate the ADL related time-frequency amplitudes from noisy amplitudes. Threshold parameters are mentioned in Section 6.4 under Parameter Initialization.

6.3.2 Event Detection from the Prescreener

Prior to extracting features for classifying a fall, *FreeFall* detects whether an ADL has occurred. For this, a PBC is used which is implemented as explained in Section 5.4.3. The PBC simply sums the energy between a predefined frequency interval (f_h and f_l) in the spectrogram at time instant t and if the summation exceeds the estimated noise energy sum within f_h and f_l, this can be considered as an event. The noise energy is estimated from the maximum noise energy above $180Hz$ where no human induced time-frequency amplitudes exist. A human induced event occurs typically within frequencies ranging from $5 - 180Hz$.

The PBC frequency levels are set as $f_l = 5Hz$ & $f_h = 25Hz$ to detect events in frequency levels common to both falls and ADLs. A range of $f_l = 5Hz$ & $f_h = 25Hz$ is used because the purpose of the prescreener is to detect ADLs, not just the falls, once the ADLs are detected, the features can be fed to the classifier to properly distinguish falls. One might suggest that setting f_h above $25Hz$ may filter more ADLs, however, the highest frequency of a fall may vary in a wide range ($30Hz - 180Hz$) depending on the speed or height of the fall. Therefore, such a frequency range is chosen such that all the falls are available for classification.

6.3.3 Activity Segmentation

From the spectrogram, velocity profiles during a detected event are extracted based on the percentile method [26] using the 95% percentile. The percentile method provides an Instantaneous Frequency (IF) for a particular ADL at a given time instant and it is correlated with the velocity of the body. The extracted IF from the percentile method is merely a measure of combined speeds of different body parts. This is, however, sufficient for segmenting the ADLs because falling is a full body motion where the torso movement has a higher impact on the instantaneous frequency of the spectrogram than other body parts.
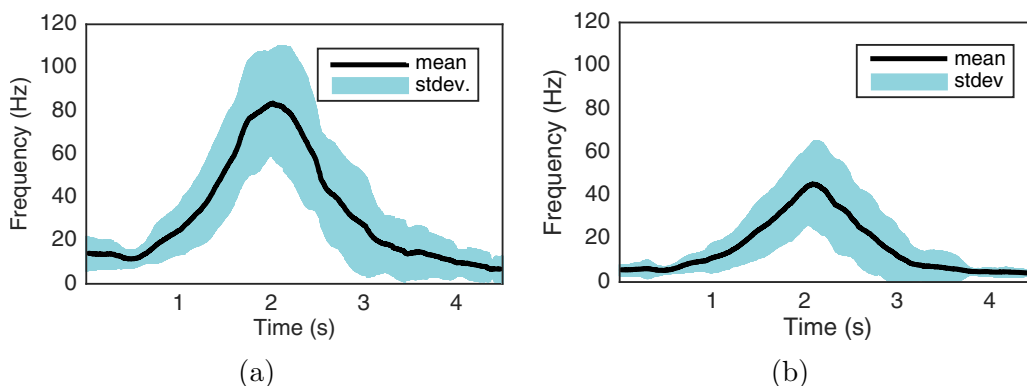
Figure 6.5: 6.5a) Frequency/velocity curve of middle ranked principal components. 6.5b) Frequency/velocity curve of a denoised subcarrier.

Since the principal components are selected from PCA such that high frequency elements have a higher impact in the spectrogram, the percentile method yields higher instantaneous frequencies than from the raw subcarriers. Figure 6.4 further illustrates that the amplitude cross section of a spectrogram constructed from this method has high amplitudes in higher frequencies than the cross section from a spectrogram of a raw subcarrier. Figure 6.5 illustrates the velocity profiles extracted from this method and a raw subcarrier for 35 falls. At the highest frequency, this has a maximum frequency difference of $\approx 30 Hz$ compared to the frequency profile extracted from the raw subcarrier.

Once an event is detected using the Power Burst Curve, the spectrogram is segmented into multiple ADLs using the velocity curve. Figure 6.6 illustrates the segmenting procedure. The velocity curve is smoothened using a local regression method, i.e. the LOcally Weighted Scatterplot Smoothing (LOWESS) method [133] using a 1st degree polynomial model and smoothing parameter $f = 0.1$ where $0 < f < 1$ before detecting peaks and troughs.

The main idea behind LOWESS is that any function can be approximated in a small neighborhood by a low-order polynomial. Smoothing parameter determines the size of the neighborhood required to fit a local polynomial. 1st degree polynomial LOWESS is a weighted linear least squares fit of the points in the local neighborhood for producing a smooth and accurate curve. The main reason for selecting LOWESS smoothing is that it gives higher priority to local points,
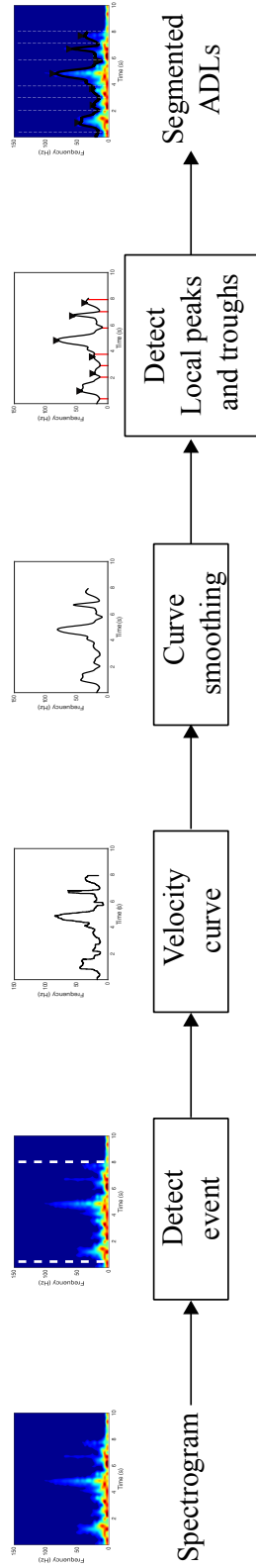
Figure 6.6: Activity segmentation flow.

(a) Event duration of a fall.

(b) Velocity curve.

(c) Cumulative displacement curve.

(d) Acceleration curve.
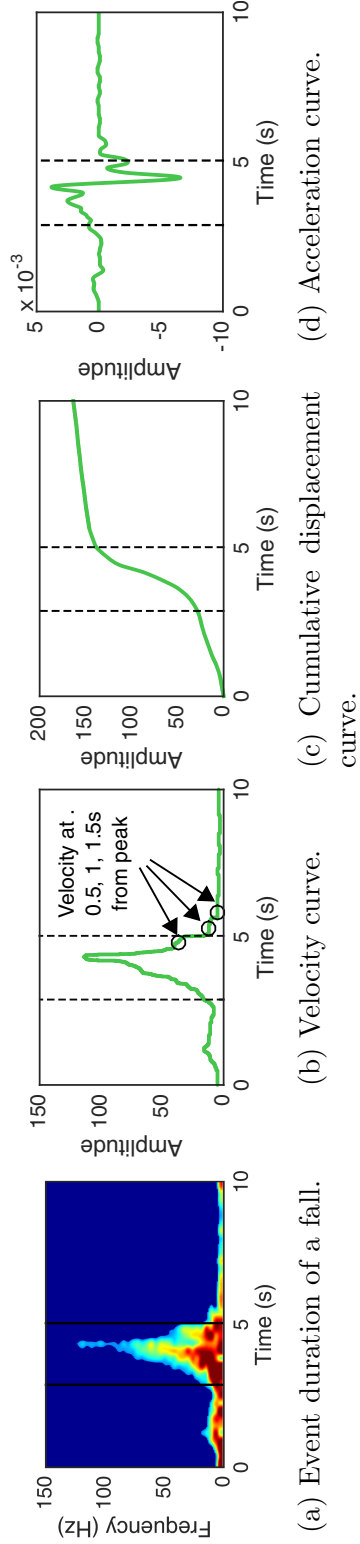
Figure 6.7: 6.7b) Velocity, 6.7c) displacement and 6.7d) acceleration of the fall in Fig. 6.7a.

therefore, the smoothed signal is not lagging the raw velocity curve. Smoothing prevents detecting multiple peaks or troughs closer to each other ($< 1s$) due to the coarse nature of the velocity curve.

Once the velocity curve is sanitized, peaks and troughs are detected in that curve. As already mentioned in Section 6.2 an ADL is located between two troughs. Based on this observation, the ADL can be isolated. Finally, features are extracted from the isolated ADL from the spectrogram.

6.3.4   Feature Extraction

Here, only the features that correspond to the understanding of a fall to differentiate from other ADLs are extracted. The velocity curve of CSI can be affected by several external factors such as position of the person [102; 120], CSI signal attenuation, and frequency components produced by different body parts, therefore the velocity curve for a single antenna pair may not provide the ground truth velocity of a fall. Therefore, to increase the accuracy two types of features (spectrogram related and velocity related) that are universal to all types of falls are extracted. To select these two types of features, a sequential forward feature selection algorithm [134] is used where the selection criterion is highest accuracy when trained and tested on two data sets from two different environments. This helps in selecting features that are resilient to changes in environments.

6.3.4.1   Spectrogram related features

■ Spectral entropy (SE) [76] *of the ADL*:

$$SE = -\sum_{n=1}^{N}\sum_{k=1}^{K} S[n,k] log_2(S[n,k]) \tag{6.1}$$

where $S[n,k]$ is the amplitude at time n and frequency k. This measures the uniformity of the time frequency distribution and a higher value indicates the energy in the spectrogram is uniformly distributed while a lower value indicates the energy is concentrated in different regions. Therefore, spectral entropy

helps to differentiate high randomness due to sudden falls from other ADLs. The spectral entropy of the whole ADL is measured up to 180 Hz and also in 10 Hz frequency steps between 40 - 180 Hz.

- Spectral entropy before and after the peak point of the ADL: During a fall it is noticable that there is power attenuation in CSI which is translated to high energy in low frequency regions (5 - 20 Hz). After the impact on the floor this attenuation diminishes as the subject is on the ground. Therefore the ADL is separated into two parts from the peak point of the velocity curve and extract spectral entropy for those two regions.

- Spectral flux (SF) [76] *of the ADL*:

$$SF = \sum_{n=1}^{N-1} \sum_{k=1}^{M-1} |S[n+1, k+q] - S[n, k]| \tag{6.2}$$

This measures the change rate of spectral components in the spectrogram within the ADL. As the fall event occurs suddenly, the spectral components have a high rate of change.

6.3.4.2 Velocity related features

- Peak velocity of the ADL: The peak point of the velocity curve is measured because it coincides with the velocity at the impact point of the body on the floor. This is shown in Fig. 6.7b

- Velocity after 0.5 s, 1 and 1.5 s from peak velocity: Low values for these velocities signify that the fall event has ended. This is further illustrated in Figure 6.7b. Velocities at 0.5, 1 and 1.5 s from the highest velocity were selected assuming a fall lasts between 2 - 3 s. Velocities before the peak velocity are not extracted because a fall can be a continuation pattern of the previous ADL, e.g. slipping while walking.

- Total displacement during the ADL: This is derived by integration of the velocity curve during the detected interval of the ADL as shown in Fig. 6.7c. This provides a measure of total distance the body traveled during the fall.

- Peak to peak acceleration: As the body accelerates before hitting the ground and there is a deceleration after hitting, the difference between the highest acceleration and highest deceleration of the body is measured during the event. This is shown in Fig 6.7d.

- Duration of the ADL: A fall typically lasts for 2 - 3 s, therefore, this feature is important to distinguish falls from other longer or sudden events in time frequency. Additionally the multiplication of ADL duration and the peak velocity is derived.

6.3.5 Classification

Binary classification was selected to identify a fall from other ADLs because the intended application only requires distinguishing a fall from another ADL, therefore, the selected features are specially suited for distinguishing a fall from an ADL. High variance in classification accuracy can be observed as a consequence of two reasons: *a)* low sample size of the data sets and *b)* class imbalance between falls and ADLs. As a solution for high variance in classification accuracy, synthetic samples were generated from the data using SMOTE [130]. As the classifier performance is dependent on the type of data, several classifiers (logistic regression, naive Bayes, SVM and random forest) were compared and random forest classifier was selected because it yielded a higher accuracy than logistic regression and naive Bayes and the results were consistent than SVM when tested on the same dataset multiple times. Unless otherwise mentioned, for training and testing on the same dataset, each dataset is divided as 70% for training and 30 % for testing.

Once the ADLs are classified as falls, to further reduce false alarms, *FreeFall* waits for a 60 s duration before alerting the caregivers. The duration is selected according to [64]. During this period, *FreeFall* observes whether another ADL occurs using it's prescreener. If the classifier detects a fall and within the next 60 s the prescreener detects another ADL, *FreeFall* classifies it as a non-fall because detection of an ADL after the fall signifies that the person is conscious.

6.4 Evaluation Setup

The hardware and software setup of *FreeFall* is implemented as mentioned in Section 5.6.2.

6.4.1 Execution

The performance of *FreeFall* is evaluated with two types of data sets. (i) an existing data set from FallDeFi (*Dataset I*) which was collected in five different environments (corridor, bedroom, kitchen, bathroom, lab) each having a single radio link. Figure 6.9 illustrates the experimental environments. (ii) a new data set *Dataset II* collected in three environments (kitchen, break room and seminar room) each having three radio links. Figure 6.10 illustrates the experimental environments. Table 6.1 illustrates the differences in each environment and the number of falls and ADLs conducted in them. The data in both data sets was collected from falls and other activities carried out by three volunteers. During the data collection period for each ADL, the transmitter sends 100 Byte packets at a data rate of 800 kbps. Therefore, during each $t_{int} = 10s$ 10000 packets are collected for each activity, corresponding to 30×10000 CSI values for each antenna pair. All experiments are performed in the 5.2 GHz band because (i) it has less radio frequency interference than in 2.4 GHz band and (ii) according to Doppler equation, the frequency components in activities are higher than in the 2.4 GHz band.

6.4.2 Parameter Initialization

FreeFall depends on input parameters at four stages: *wavelet de-noising, spectrogram processing, pre-screener* and *classification*. De-noising and classification parameters were selected once for all the environments while spectrogram processing and pre-screener parameters change dynamically. The DWT-based denoising strategy that was developed in FallDeFi is implemented in FreeFall as it is. The parameter initialization in the denoising step is mentioned in Section 5.6.4.1.

The threshold $\hat{N}_{th}$ used for spectrogram segmentation is estimated from the maximum spectrogram amplitude in frequencies above 180 Hz. As human induced

Table 6.1: Collection of two data sets, Dataset I in five different environments and Dataset II in three different environments and the characteristics of each environment.

| | | Dataset I | | | | | Dataset II | | |
		Corridor	Bedrooms	Kitchen	Bathroom	Lab	Kitchen	Break room	Seminar room
# Falls		82	51	69	75	49	50	100	120
# ADLs		243	218	355	324	232	131	148	232
Environment		uncluttered	cluttered	uncluttered	uncluttered	cluttered	uncluttered	uncluttered	cluttered
Area size (m^2)		9.9	16.12	18.9	11.6	33.58	18.9	64.9	66.49
Link type		LoS	nLoS, through-wall	LoS	nLoS, through-wall	nLoS	LoS	LoS	LoS
Link length (m)	TX-RX1	9	5	4	5	7	3.5	7	4
	TX-RX2	-	-	-	-	-	4	6.5	9
	TX-RX3	-	-	-	-	-	4.5	8	10

events occur typically within frequencies 5 - 180 Hz [120], amplitudes beyond 180 Hz frequency can be assumed as noise. Since the pre-screener threshold (PBC_{th}) depends on this threshold, the event detection step is also impacted by this value.

The number of trees in the classifier was set to 50. This choice was made by measuring the out-of-bag classification error with increasing number of trees. The minimum number of trees that corresponds to a minimum in classification error was selected here. Minimum leaf size (number of observations in the last node) was set to five to avoid over-fitting to a particular environment. The prior probabilities of each class and the cost of misclassification of a fall were adjusted to the ratio between the number of falls and non-falls in the training set in order to yield a high penalty for misclassifying a fall.

6.4.3 Types of Activities

Falls and ADLs in both datasets were performed by three volunteers aged between 27-30 years old with different physiques. The conducted fall types included trip, slip, lose consciousness and lose balance. These falls can occur in four different directions with respect to the body: forward, backward, lateral or as on-position. As each person may imitate the falls in a unique way, the above variations in falls were considered to avoid over-fitting to a particular falling pattern of a single person. As for ADLs, only the ones that can be easily mistaken as a fall are considered. These include standing up from a chair/floor, sitting down on chair/floor, bending down and picking up an object, jumping, walking and pausing, no event and opening or closing doors. In all environments the experiments were performed, there was either only one volunteer in the area of interest or there were two of them. During the latter scenario, when one person performs the scripted ADLs, the other person is either motionless or performs regular activities. More than two volunteers were not considered in the experiments because many falls of elderly people go unnoticed when they live independently which is the worst case scenario.

6.4.4 Evaluation Metrics.

Note that the evaluation metrics used in this work are slightly different from the metrics used in FallDeFi in Section 5.6.6. The metrics used in here are as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{6.3}$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6.4}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{6.5}$$

where TP = true positives, FP = false positives, TN = true negatives and FN = false negatives. True positives are falls that are manually annotated and correctly detected by the classifier, false positives are non-falls that are detected as falls, true negatives are non-falls that are correctly detected as non-falls and false negatives are falls that are misclassified as non-falls. Ideally, accuracy and TPR should be close to 100% and FPR should be close to 0%. The true positive rate provides the fraction of falls that are correctly identified among actual falls. The evaluation metrics used in Section 5.6.6 for are accuracy, f1 score, precision and recall. Recall is exactly the same as true positive rate and f1 score is a derived metric from precision and recall. Therefore, precision and false positive rate are the main differences in the metrics used here. Precision measures the fraction of correctly identified falls from all detected falls which together with accuracy is an indirect measure of false alarms, while FPR identifies the fraction of relevant non-falls among all the detected non-falls, therefore it is a direct measure of false alarms.

6.5 Results

In this section a performance analysis of *FreeFall* is provided using the collected data. *Dataset II* is used to study how a fall is impacted by different access points, both from the perspective of the location of the fall and the access point location as well as the length of the radio link. This provides an indication on access

point locations in order to detect the falls effectively, this is discussed in detail in Section 6.5.1. *Dataset I* is used to evaluate the robustness of *FreeFall* in four different environments in LoS, non LoS, cluttered and through-wall conditions. It is also evaluated how *FreeFall* adjusts to changes in the environment and finally the approach is compared with the state of the art under the same conditions.

6.5.1 Performance depending on Room Size, Access Point Placement and Location of the Fall

Figure 6.10 illustrates the locations in kitchen, break room and seminar room where the falls for *Dataset II* took place. A mattress is placed in the numbered positions and performed 10 falls were performed in each location. As shown in Table 6.1, a total of 50, 100 and 120 falls were performed in the respective environments and those falls were captured by three links (AP0-AP1, AP0-AP2 and AP0-AP3) in each room. Note that AP0 acts only as the transmitter and AP1, AP2 and AP3 act as receivers.

(a) Apartment

✕ Falls and other activities   ✕ Other activities

(c) Bathroom

(b) WSN Lab

Figure 6.9: The floor plans of each environment in Dataset I. [©David Rojas]

(a) Kitchen.

(b) Break room.

(c) Seminar room.
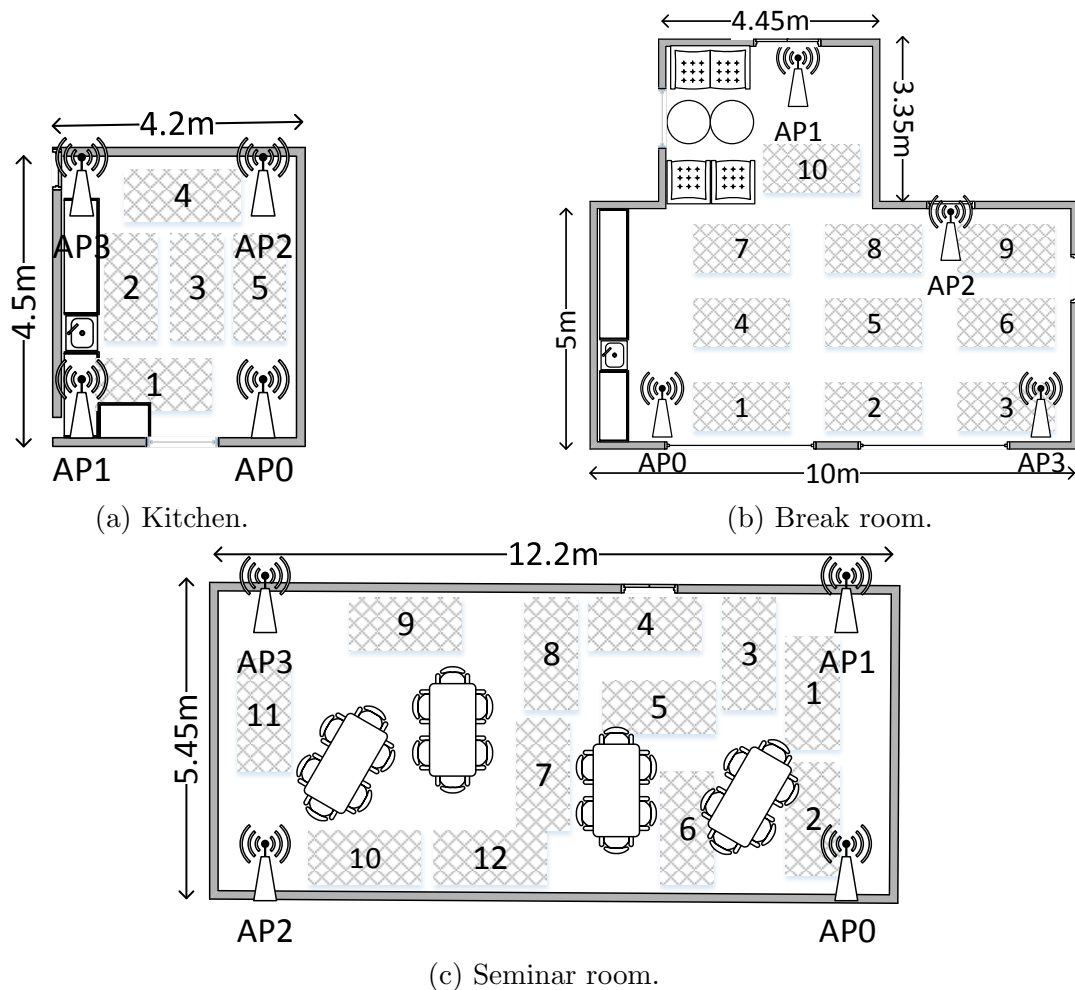
Figure 6.10: The floor plans of each environment in Dataset II. Each environment has 4 access points, one transmitter and three receivers. The experiment positions are numbered.

Event detection and fall detection results of the three links in each environment are shown in Table 6.2. These two results are separately reported in order to analyze the performance of both pre-screener and the fall detection system individually. The pre-screener results indicate the percentage of falls correctly identified as an event out of the total number of actual falls performed. Accuracy, TPR and FPR are the metrics for the performance of the fall detection system. As the number of ADLs is higher than the number of falls in each room, the FPRs have a higher impact on the accuracy. Out of the three rooms, all the

three links in the kitchen have accuracies and TPRs over 90%, i.e., accuracy of 92.31%, 96% and TPR of 99.44% and 91.11%, 97.96% and 98%. Comparatively, the other two environments have only one link that has over 90% accuracy and TPR, i.e. link AP0-AP3 in both break room and seminar room have accuracies over 90%. This result is quite expected as the area size of the kitchen is $\approx$ 3.5 times smaller than the other two rooms and each link on its own has enough coverage to detect falls in the entire area of the kitchen. Consequently, AP0-AP3 link in both the rooms has the highest coverage for falls because it lies in the diagonal.

Link AP0-AP3 in all the three rooms has the best event detection (100%, 99% and 100%) and fall detection accuracy (99.44%, 95.58% and 94.75%) and AP0-AP1 has the worst. One common characteristic of the best performing links is that they are the longest in each room e.g. AP0-AP3 in the three rooms is 4.5m, 8m and 10m long. When a link is lengthy, the coverage of the room increases as the FFZ widens. This allows a wider coverage for the positions where the experiments were performed and results in a higher detection of falls. The opposite of this also holds as the shortest links (AP0-AP1) tend to have the worst results, e.g. 92.31%, 85.89% and 88.53% accuracy in the kitchen, break room and seminar room respectively.

The detection of falls is further analyzed here in individual positions compared to the radio link. Figures 6.11 and 6.12 show the detection of events and falls in each position of the three rooms. At position 4 in the kitchen, AP0-AP1, which is the worst position for the worst link, has a 70% event detection and a 90% fall detection accuracy which corresponds to only $\approx$ 63% of falls that are correctly detected. In the break room, almost all positions for link AP0-AP1 are affected except positions 1 and 7. The worst result is at positions 5 and 9 with $\approx$ 60% fall detection accuracy. In the seminar room all positions, except the positions closest to the link (positions 1, 2 and 3), are affected. One clear reason for this behavior is the coverage of the FFZ. As an example, the semi-minor axis of the FFZ having an elliptical coverage model [96] for link AP0-AP1 in the kitchen is $\approx$ $3m$ which does not reach position 4 and in the seminar room it is $\approx$ $3.5m$, which does not extend beyond position 3. In the break room, the line of sight is obstructed by furniture in the middle and the concrete wall on the right hand

6.5 Results

Table 6.2: Event detection and fall detection results for the three environments in Dataset II.

%	Kichen			Break room			Seminar room		
	AP1	AP2	AP3	AP1	AP2	AP3	AP1	AP2	AP3
Event detection	92	100	100	96	98	99	94.17	95	100
Accuracy	92.31	96	**99.44**	85.89	88.94	**95.58**	88.53	88.34	**94.75**
TPR	91.11	97.96	**98**	78.13	81.63	**94.95**	77.67	80.53	**98.36**
FPR	7.20	4.64	**0**	9.21	5.04	**4.05**	4.19	7.51	**6.65**

side limits most of the fall detections on that side.

The above results indicate that room size, link length, the link's position in the room and the position of the fall relative to the link all have an impact on the overall fall detection accuracy of *FreeFall*. It is also evident that by improving the WiFi coverage, especially the FFZ, better results can be obtained. Furthermore, by optimizing the positions of the access points, the coverage can be improved. As an example, for all three environments, the best performing links (AP0-AP3) achieve $\geq 94.75\%$ accuracy, $\geq 94.95\%$ TPR and $\leq 6.65\%$ FPR.

6.5.2 Robustness of the Features to Changes in Conditions

Dataset I is used to evaluate how the extracted features perform when *FreeFall* is trained with data from one environment and tested on another. The collected fall and ADL data from *Lab* are used for training and the data from *Corridor*, *Bedrooms*, *Kitchen* and *Bathroom* are used for testing. Note that *Bedrooms* have a through wall link (single wall with decoupled and insulated plaster partition) and *Bathroom* has a link penetrating two walls with plaster partitions. Other major differences in the links and environments are shown in Table 6.1. When trained and tested for the same environment, *FreeFall* achieves an average accuracy of 94.7%, TPR of 96.33% and FPR of 5.63%. When trained with Lab data and then testing on data from other rooms yields a decrease of 1.82% in accuracy, 8% decrease in TPR and 9% increase in FPRs. This implies that the classifier and the extracted features in Section 6.3.4 are able to classify ≈ 8.9 out of 10 falls (88.59% TPRs) on average even when the trained environment changes.

Average performance indicates only the general tendencies of the features across all conditions when they are trained in the *Lab* and tested in other places. Therefore, the robustness of the features are also analyzed among individual environments and then the links are divided into LoS and non LoS. Table 6.3 illustrates those results. *Kitchen*, *Bedroom* and *Bathroom* have over 90% accuracy and *Corridor* has 88.5%. Again, the FPRs has a higher impact on the accuracy than TPRs because the number of ADLs are greater than twice the number of falls. The extracted features perform best in terms of FPR in Kitchen (100%)
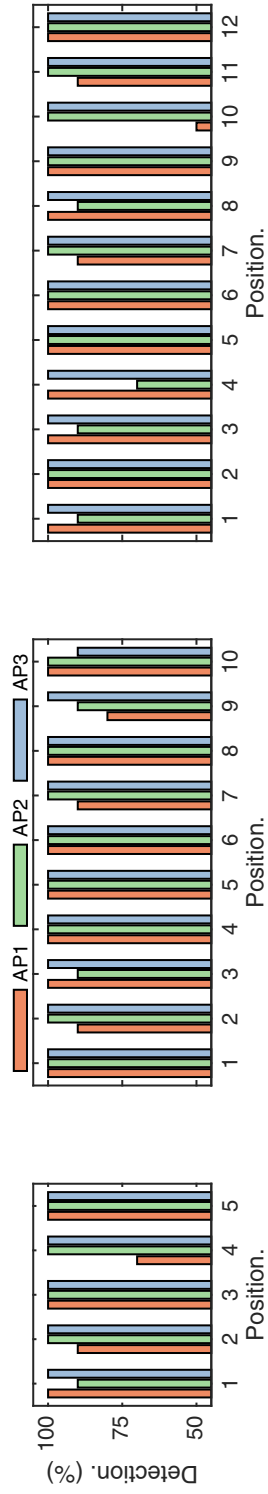
Figure 6.11: Prescreener performance of individual positions.
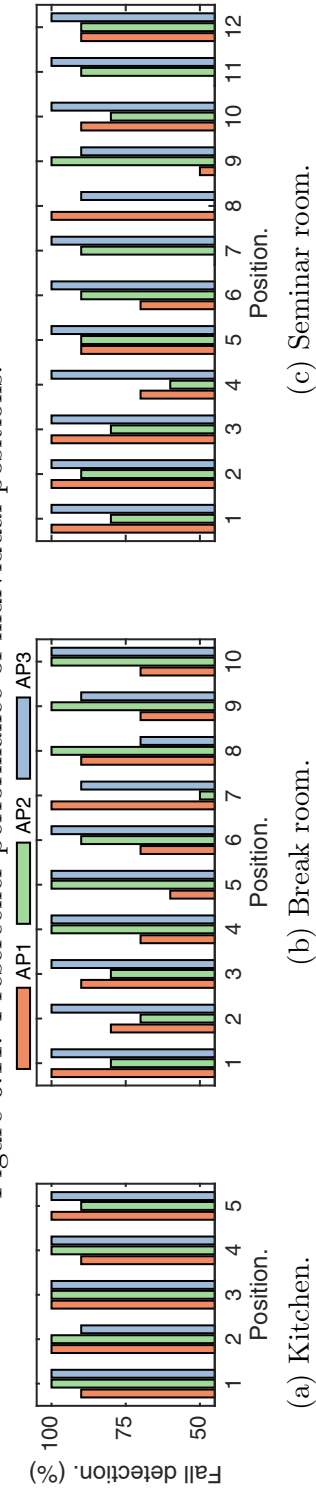
Figure 6.12: Fall detection performance of individual positions.

and the worst performance is in the Corridor (79.54%). The Corridor being a narrow passage having a width of $1.1m$ and the access points placed $1m$ above the ground tends to deteriorate the detection performance due to shadowing of the line of sight.

In Table 6.3 the comparison is across LoS and non LoS conditions. Corridor and Kitchen are grouped into LoS and Bedroom and Bathroom are grouped into nLoS. The average results in the table indicate that the non LoS conditions have lesser false alarms than LoS as FPRs are $< 6.43\%$. The TPRs are $> 3\%$, in LoS indicating a better detection of falls. As the FPRs have a higher impact on the accuracy than TPRs, due to high number of ADLs than falls, the average accuracy in nLoS is higher than in LoS. Recall that *FreeFall* is trained using features extracted from *Lab* environment for these results. *Lab* being a nLoS environment seems to produce more false positives and true positives for the LoS conditions than non LoS.

6.5.3 Comparison with the State of the Art

Table 6.4 compares the performance of *FreeFall* with three other systems using *Dataset I*, FallDeFi presented in Chapter 5, CARM [26], a CSI-based activity recognition scheme, and RTFall [100], a fall detection scheme using the CSI phase. As already mentioned in Section 5.7.4 CARM was selected because FallDeFi is inspired by the signal processing techniques used in CARM and it is the most comparable activity recognition scheme to FallDeFi in literature. Even though there are two other CSI-based fall detection schemes in the literature, RT-Fall was selected specifically because it builds upon the features used by those two approaches. Comparison results are shown in Table 6.4, *FreeFall* has the highest accuracy, TPRs and the lowest FPRs when it is deployed in the trained environment as well as when environment conditions change. *FreeFall* has an increase of 6.12% in accuracy, 2.12% increase in TPRs and 62.33% decrease in FPRs compared to FallDeFi, the next best system when trained and tested in the same environment. When the trained environment is different from the tested, the improvement in accuracy is 16% and TPR is 8.6%. The most notable improvement is the decrease of FPR by 71.52%. The main differences in *FreeFall* from FallDeFi

138

that contributed to the improvements in TPRs and accuracy are the activity segmentation to identify the boundaries between two ADLs, amplitude sanitation and the additional features. Specially, the 30 s waiting period for another ADL to occur helped in reducing the FPR significantly.

6.5.4 Comparison with Accelerometer based Fall Detection

Finally, a state of the art accelerometer based fall detection scheme [63] compares results of Kitchen and Break room of *Dataset II*. For experiments, an accelerometer (Bosch, maximum range: 39.227 ms^{-2}, resolution: 0.002 ms^{-2}, sampling rate: 200 Hz) in a smartphone is used. Raw accelerometer data were collected from the smartphone and signal processing steps are implemented as published extracting features from the accelerometer only which was worn at the waist (attached to the belt). Once the accelerometer readings are acquired for a particular activity, they are downsampled from 200 samples/s to 50 samples/s. Then a 2nd order Butterworth filter with passband 0.15 - 15 Hz removes noise and reduces the drift. Since the accelerometer readings are provided in three axes, the resultant root sum of squares (RSS(t)) is obtained for feature extraction where

$$RSS(t) = \sqrt{a_x^2(t) + a_y^2(t) + a_z^2(t)} \tag{6.6}$$

where a_x, a_y and a_z are the accelerations along x, y and z axes. Then the following features of the RSS vector are extracted:

a) Upper Peak Value (UPV) - the maximum value from the RSS signal,

b) Lower Peak Value (LPV) - the minimum value window of 2 *s* before the UPV,

c) Lead-time window between the LPV to the UPV,

d) Maximum waist segment posture angle $max(\theta(t))$, where $\theta(t)$

$$\theta(t) = cos^{-1}(\frac{\vec{a}(t) \times \vec{g}}{|\vec{a}(t)| \times |\vec{g}|}) \tag{6.7}$$

139

Table 6.3: Performance in each environment of Dataset I and performance in LoS and nLoS conditions.

%	LoS			nLoS		
	Corridor	Kitchen	Total	Bedroom	Bathroom	Total
Acc.	88.50	92.58	90.92	92.19	97.2	94.94
TPR	79.54	100	91.59	82.35	88.57	88.57
FPR	9.87	8.73	9.2	5.5	0.92	2.77

Table 6.4: Comparison of FreeFall with the state of the art for Dataset I.

%	Train and test in the same place				Train and test in different places			
	FreeFall	FallDeFi	CARM	RTFall	FreeFall	FallDeFi	CARM	RTFall
Accuracy	94.7	88.9	77.85	80.96	92.98	80.10	64.26	51.14
TPRs	96.33	94.33	80.32	85.20	88.59	81.56	53.07	67.82
FPRs	5.63	14.92	24.63	24	6.14	21.56	24.55	65.91

, $\vec{a}(t)$ is the accelerometer output vector at time t and $\vec{g}$ is the reference gravity vector.

e) Fall angle time - time at which the maximum waist segment posture angle occurs from UPV in a 3 s window

f) Maximum vertical acceleration - $max(RSSI(t) - g)$,

g) Maximum vertical velocity and - $max(\int\limits_{t} RSSI(t) - g\,dt)$

h) Maximum vertical displacement - $max(\iint\limits_{t} RSSI(t) - g\,dt\,dt)$

Once the features are extracted from the detected events, the random forest classifier is used with 50 trees, prior probability and cost of misclassification of [1, 0.8] for falls and non-falls. From the ADLs classified in those two environments, a 100% accuracy and TPRs and 0% FPRs are obtained. For the same environments *FreeFall* obtained an average accuracy of 97.4% TPR of 95.97% and FPR of 2.3%.

6.6 Discussion and Conclusion

This chapter presented several enhancements to the fall detection system developed in Chapter 5 that include amplitude sanitation and principal component selection for spectrogram construction and processing, activity segmentation for meaningful feature extraction and more importantly improved the accuracy and drastically reduced the FPRs by extraction of novel features. Moreover, it was shown that the extracted features are resilient to changes in environments. Further experiments were performed to identify the influence of fall position, access point position and room size on fall detection accuracy. Through the hardware implementation, experimentation and evaluation in multiple environments, it is shown that the fall detection system performs better than existing CSI-based fall detection systems. A 94.7% average accuracy is achieved for a system trained and tested in the same environment while the average accuracy is closer to 93% when tested in a completely different environment.

FreeFall has a dramatic improvement on the current state of the art, yet, the system is a research prototype and requires further analysis and improvements before it can finally be deployed in the real world to save the lives of elderly people. As an example, Section 6.5.4 demonstrates that accelerometer based fall detection is still more accurate than *FreeFall*. As the results show, specially when the falls are away from the FFZ, depending on the link type, *FreeFall* may fail to detect a fall. A wearable system like an accelerometer does not have location dependent failures. However, forgetting to wear such a device is the main drawback of using accelerometer based fall detection devices. Therefore, it is envisaged that when using these two systems together, the drawbacks of both systems can be minimized and achieve a far better accuracy.

One particular area where *FreeFall* requires improvement is increasing the accuracy in confined places like corridors. As an example in Table 6.3 *Corridor* has the lowest accuracy which is 88.5%. The reason can be conjectured as the shadowing effect that causes the accuracy drop. However, further experimentation is required in such places, as an example, by varying heights of the transceivers to identify the optimum height.

Another area that requires further attention is the experimentation on falls that occur in low heights and low speeds, e.g. falling from a chair. These types of falls may not be as life threatening as falls that create high impact on the floor, however, the system should be fine tuned to detect and classify such events in situations where the human subjects require external assistance.

Fall data was collected from volunteers aged between 27-30 years as it is not really possible to subject elderly people to large number of falls in order to gather the required experimental data. This most likely requires adaptation of the system parameters for elderly people that fall in real world situations. It requires further analysis on adaptation mechanisms of the system to such scenarios. As an example classifier prior probabilities and costs can be tuned to be more sensitive to falls and ADLs having slow or fast movements depending on the human subject's character.

In the current implementation of *FreeFall*, when a fall is detected, it is considered a non-fall if an ADL is detected within 30 s. This can be further improved by classifying what type of ADL is detected after the fall in order for the system

to make better decisions on informing the caregivers. Additionally, the system can incorporate a push button as assistance to the human subjects if they require help of outside once they get up after a fall.

Chapter 7

Conclusions and Future Work

7.1 Summary

This thesis presented research on a device-free, low cost, indoor fall detection system using existing WiFi infrastructure. As main takeaways, this thesis provided state of the art signal processing methods, toolchains and features for human presence detection, event detection and fall detection using WiFi CSI. Detecting binary occupancy in the observation environment is performed before detecting falls. Once the human presence is detected, the fall detection system first detects human induced events, segments those events into human activities and infers falls from those activities. The developed modules, signal processing methods and features in this work for the mentioned functionalities can act as a stepping stone for future researchers working towards developing research modules and applications for industry needs in fall detection, using commercial-off-the-shelf WiFi components. In future fall detection applications, miniaturisation and the low-cost of commercial WiFi chips can help to drive down deployment costs of fall detection systems, which increases their chances of expansion in an area that is impacted by reliability and cost over-run issues.

7.2 Research Contributions

This thesis provided three main contributions.

a) Detecting binary occupancy or whether there is a person in the observation environment is a precursor to detecting falls. As the first contribution of this thesis, a novel time domain human presence detection scheme was developed using a non-linear dimensionality reduction method through the understanding of the behavior of CSI under human motion. Similar to existing CSI based human presence detection schemes, this work performs threshold-based detection using eigendecomposition. Unlike previous work that uses linear techniques for this, the presented solution is derived through exploiting correlations among subcarriers using a nonlinear method. The system was implemented using commercial WiFi devices and experiments were performed in real world environments.

b) The next step is the design and development of a novel fall detection system which is called "FallDeFi" and this is a culmination of four different contributions.

- A novel CSI noise filtering and dimensionality reduction approach to facilitate feature extraction in high frequencies.

- Extract novel features in the time-frequency domain to detect falls and a rigorous feature selection procedure to filter the features that are resilient to changes in the deployed environment.

- Development of a novel time-frequency based pre-screener for event detection prior to detecting falls and using its outputs, such as event duration, as features in fall detection.

- Evaluation of the performance of the proposed system with real world experiments using commodity hardware. The performance was evaluated through experiments using human subjects in four different environments, having different link characteristics (LoS, none LoS, through-wall) and up to two persons using the data collected from commodity WiFi devices. The results demonstrate improvements in accuracy compared to two state of the art fall detection systems.

c) Even though FallDeFi improves on the state of the art, the main challenge of this system was identified as improving the accuracy and reducing the

145

number of false alarms for real world deployment. Two main enhancements to the fall detection architecture of FallDeFi were made to improve detection performance. Further experiments were carried out to analyze the performance of the system and the final system is called "FreeFall". The three main contributions of FreeFall are as follows.

- FallDeFi initially detects a human event through a prescreener and then classifies the activity in that event assuming a single activity occurs during the detected event duration. This can lead to misclassification because the extracted features contain information from a mixture of events. This problem is addressed by an activity segmentation procedure that uses human velocity profiles.

  A combination of in-band noise filtering and a novel principal component selection algorithm removes the artifacts in low frequencies caused by multipath propagation, line of sight attenuation and other interference terms due to signal processing. This helps to obtain precise human velocity profiles from the spectrogram.

- Additional features were extracted from the segmented activities to improve the accuracy and true positive rate and feature selection was used to enhance resilience to environment changes.

- As the final contribution, experiments were carried out in three different environments to analyze the impact on fall detection from falling position, distance from link and the impact of the size of the room. The performance of FreeFall is compared with FallDeFi and two other state of the art fall detection systems and FreeFall achieved higher accuracy, true positives rates and lower false positives rates compared to the other three systems.

7.3 Future Research

Although the final system improves on the current state of the art, this nevertheless is a research prototype implementation and requires further improvements before it can finally be deployed for the intended application of saving lives of

elderly people. This work can be further enhanced by the fusion of location information of the targets. As mentioned in Section 5.7.2.3, the target position has an impact on distinguishing falls that occur far away from the link. This is also confirmed by radar based fall detection research [135], where occupant range information has increased the detection of falls. Occupant range information can be obtained from the same WiFi infrastructure used in FallDeFi and FreeFall in a device-free manner [43; 44].

Falls can be further distinguished from other activities by capturing the movement direction obtained through Doppler signatures [95]. For this, it requires precise phase information of CSI with further refinement of the signal. Currently, the system uses only CSI amplitude measurements from one transceiver pair dedicated to monitor falls in a particular area.

Prediction of the likelihood of a fall or the risk of falling of a human subject by the quantification of his or her walking speed or gait velocity is another research direction. Previous research has indicated that gait velocity of a person is an important measure of falling risk [136]. It has been shown to be a risk factor for disability, cognitive impairment, falls, and mortality [137]. In fact, gait velocity is considered such a reliable and sensitive measure that it was recommended to be the sixth vital sign [138] along with other vital signs such as body temperature, pulse rate, respiration rate and blood pressure. Previous research has indicated that WiFi signals can be used to measure human gait accurately [101; 139]. Therefore, by measuring the gait velocity of each person beforehand and continuous monitoring for any abnormalities of walking during everyday life, this metric can quantify the risk of falling and also influence the fall detection accuracy by affecting the likelihood function of a fall for each person.

As mentioned in Chapter 6, by fusing WiFi and accelerometer based fall detection, the drawbacks of the two systems can be mitigated. As the results show, eQQ Q specially when the falls are away from the FFZ, depending on the link type, FreeFall may fail to detect a fall. A wearable system like an accelerometer does not have location dependent failures. However, forgetting to wear such a device is the main drawback of accelerometers. Therefore, it is envisaged that when using these two systems together, the drawbacks of both systems can be minimized and achieve a far better accuracy. There is another line

of research where floor vibrations measured by accelerometers or other vibration sensors detect falls [16]. These devices operate in a smaller range than WiFi, yet, they can be deployed in places where WiFi coverage is limited. This data can be further fused to the existing WiFi measurements to reduce the dependence on falling position and enhance the accuracy.

The ideal validation of this system's performance would be a long-term deployment of a large WiFi network in a domestic environment in real time. However, this type of implementation is beyond the available time and resources of this PhD. This implies that the problems or weaknesses not identified from the current implementation could be exposed by a long-term deployment, particularly in reliability. Even though this work has potential to work in real time as mentioned in Section 5.7.4, real time operation is yet to be implemented. This requires the collection of CSI from WiFi transceivers and transferring them to a separate server for signal processing. As 'CSI Tool' has been used throughout this work for obtaining CSI, it has to be further developed to operate in real time. Another important step is testing the system on environments where elderly people live, however, it is not possible to subject elderly people to emulate the action of falls. As current research using accelerometers has demonstrated, it requires a period of 6 to 12 months to obtain a sufficient number of samples in such environments.

It is further required to analyze how the system performs in the 2.4 GHz band where multiple wireless technologies coexist. Recent research has shown that cross technology interference can impact the accuracy of an activity recognition system [140]. Fall detection being subdomain of activity recognition, interference could more likely impact the current system in this band.

Although CSI is used in all 802.11n compatible chipsets in the physical layer, it can be accessed in upper layers only in a limited number of chipset types through driver and firmware modifications e.g. Intel 5300 and Atheros 9k series. However, this constraint is not just limited to FreeFall. With the expansion of research that use CSI as a signal strength descriptor, the exposure of CSI in a wide range of products can be expected.

Bibliography

[1] Ryan Melfi, Ben Rosenblum, Bruce Nordman, and Ken Christensen. Measuring building occupancy using existing network infrastructure. In *IEEE International Green Computing Conference and Workshops*, pages 1–8, 2011. xiv, 7, 8

[2] Thomas Paul and Tokunbdo Ogunfrunmiri. Wireless lan comes of age: Understanding the ieee 802.11 n amendment. *IEEE Circuits and Systems Magazine*, 8(1):28–54, 2008. xiv, 27

[3] Yiwei Zhuo, Hongzi Zhu, Hua Xue, and Shan Chang. Perceiving accurate CSI phases with commodity WiFi devices. In *IEEE Conference on Computer Communications*, pages 1–9, 2017. xiv, 28

[4] United Nations, Department of Economic and Social Affairs, Population Division (2015). World Population Prospects: The 2015 Revision, custom data acquired via website. `https://esa.un.org/unpd/wpp/`, 2015. Accessed: 2017-04-24. 1

[5] Bruce H Alexander, Frederick P Rivara, and Marsha E Wolf. The cost and frequency of hospitalization for fall-related injuries in older adults. *American journal of public health*, 82(7):1020–1023, 1992. 1

[6] Daniel A Sterling, Judith A OConnor, and John Bonadies. Geriatric falls: injury severity is high and disproportionate to mechanism. *Journal of Trauma and Acute Care Surgery*, 50(1):116–119, 2001. 1

[7] Alice C Scheffer et al. Fear of falling: measurement strategy, prevalence,

risk factors and consequences among older persons. *Age and ageing*, 37(1): 19–24, 2008. 1

[8] Marilyn Rantz, Marjorie Skubic, Carmen Abbott, Colleen Galambos, Mihail Popescu, James Keller, Erik Stone, Jessie Back, Steven J Miller, and Gregory F Petroski. Automated in-home fall risk assessment and detection sensor system for elders. *The Gerontologist*, 55(Suppl_1):S78–S87, 2015. 1

[9] Muhammad Mubashir, Ling Shao, and Luke Seed. A survey on fall detection: Principles and approaches. *Neurocomputing*, 100:144–152, 2013. 1, 2, 15, 16

[10] Raul Igual et al. Challenges, issues and trends in fall detection systems. *Biomedical engineering online*, 12(1):66, 2013. 1

[11] Natthapon Pannurat et al. Automatic fall monitoring: a review. *Sensors*, 14(7):12900–12936, 2014. 2

[12] Moeness G Amin, Yimin D Zhang, Fauzia Ahmad, and KC Dominic Ho. Radar signal processing for elderly fall detection: The future for in-home monitoring. *IEEE Signal Processing Magazine*, 33(2):71–80, 2016. 2, 17, 18, 67, 76

[13] Christian Debes, Andreas Merentitis, Sergey Sukhanov, Maria Niessen, Nikolaos Frangiadakis, and Alexander Bauer. Monitoring activities of daily living in smart homes: Understanding human behavior. *IEEE Signal Processing Magazine*, 33(2):81–94, 2016. 2, 19

[14] C. Han et al. WiFall: Device-free fall detection by wireless networks. In *IEEE Conference on Computer Communications*, 2014. 2, 21, 23, 114

[15] Yueng Santiago Delahoz and Miguel Angel Labrador. Survey on fall detection and fall prevention using wearable and external sensors. *Sensors*, 14 (10), 2014. 2

[16] R. Ramezani, Y. Xiao, and A. Naeim. Sensing-Fi: Wi-Fi CSI and accelerometer fusion system for fall detection. In *IEEE EMBS International*

Conference on Biomedical Health Informatics, pages 402–405, 2018. 2, 23, 148

[17] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni. FIMD: Fine-grained Device-free Motion Detection. In *IEEE 18th Int. Conf. on Parallel and Distributed Systems*, 2012. 8, 9

[18] Kun Qian et al. PADS: Passive detection of moving targets with dynamic speed using PHY layer information. In *IEEE 20th Int. Conf. Parallel and Distributed Systems*, pages 1–8, 2014. 8, 9

[19] Zimu Zhou, Zheng Yang, Chenshu Wu, Longfei Shangguan, and Yunhao Liu. Omnidirectional coverage for device-free passive human detection. *IEEE Transactions on Parallel and Distributed Systems*, 25(7):1819–1829, 2014. 8

[20] Jiang Xiao, Kaishun Wu, Youwen Yi, Lu Wang, and Lionel M Ni. Pilot: Passive device-free indoor localization using channel state information. In *IEEE 33rd international conference on Distributed computing systems*, pages 236–245, 2013. 8, 12

[21] Liangyi Gong et al. An adaptive wireless passive human detection via fine-grained physical layer information. *Ad Hoc Networks*, 2015. 8

[22] Fu Xiao et al. Invisible Cloak Fails: CSI-based Passive Human Detection. In *Proceedings of the 1st ACM Workshop on Context Sensing and Activity Recognition*, 2015. 8

[23] Z. Zhou et al. Towards omnidirectional passive human detection. In *Procedings of IEEE Conference on Computer Communications*, 2013. 8

[24] Yossi Rubner and Carlo Tomasi. *Perceptual metrics for image database navigation*, volume 594. Springer Science & Business Media, 2013. 8

[25] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao. Non-invasive detection of moving and stationary human with wifi. *IEEE Journal on Selected Areas in Communications*, 33(11): 2329–2342, 2015. 9, 10

[26] Wei Wang, Alex X. Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. Understanding and Modeling of WiFi Signal Based Human Activity Recognition. In *Procedings of the 21st Annual Int. Conf. on Mobile Computing and Networking*, 2015. 9, 20, 21, 22, 24, 30, 32, 37, 48, 67, 69, 70, 72, 74, 78, 93, 103, 104, 111, 114, 115, 117, 121, 138

[27] Ruth Ravichandran, Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. Wibreathe: Estimating respiration rate using wireless signals in natural settings in the home. In *IEEE International Conference on Pervasive Computing and Communications*, 2015. 10

[28] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. Human Respiration Detection with Commodity Wifi Devices: Do User Location and Body Orientation Matter? In *Procedings of the ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing*, 2016. 10

[29] Xuyu Wang, Chao Yang, and Shiwen Mao. TensorBeat: Tensor Decomposition for Monitoring Multiperson Breathing Beats with Commodity WiFi. *ACM Trans. Intell. Syst. Technol.*, 9(1):8:1–8:27, 2017. 10, 21

[30] C. Chen, Y. Han, Y. Chen, H. Q. Lai, F. Zhang, B. Wang, and K. J. R. Liu. Tr-breath: Time-reversal breathing rate estimation and detection. *IEEE Transactions on Biomedical Engineering*, 65(3):489–501, 2018. 10

[31] Elahe Soltanaghaei, Avinash Kalyanaraman, and Kamin Whitehouse. Peripheral WiFi Vision: Exploiting Multipath Reflections for More Sensitive Human Sensing. In *Procedings of the 4th Int. Workshop on Physical Analytics*, pages 13–18, 2017. 10

[32] W. Xi, J. Zhao, X. Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang. Electronic frog eye: Counting crowd using WiFi. In *IEEE Conference on Computer Communications*, pages 361–369, April 2014. 10

[33] Saandeep Depatla, Arjun Muralidharan, and Yasamin Mostofi. Occupancy estimation using only WiFi power measurements. *IEEE Journal on Selected Areas in Communications*, 33(7):1381–1393, 2015. 10

[34] Han Zou, Yuxun Zhou, Jianfei Yang, Weixi Gu, Lihua Xie, and Costas Spanos. Freecount: Device-free crowd counting with commodity WiFi. In *IEEE Global Communications Conference*, pages 1–6, 2017. 10

[35] T. Teixeira, G. Dublon, and A. Savvides. A survey of human-sensing: Methods for detecting presence, count, location, track, and identity. Technical report, ENALAB, 2010. 11

[36] Joey Wilson and Neal Patwari. Radio tomographic imaging with wireless networks. *IEEE Transactions on Mobile Computing*, 9(5):621–632, 2010. 11

[37] Frederic Thouin, Santosh Nannuru, and Mark Coates. Multi-target tracking for measurement models with additive contributions. In *Proceedings of the 14th International Conference on Information Fusion*, pages 1–8. IEEE, 2011. 11

[38] Santosh Nannuru, Yunpeng Li, Yan Zeng, Mark Coates, and Bo Yang. Radio-frequency tomography for passive indoor multitarget tracking. *IEEE Transactions on Mobile Computing*, 12(12):2322–2333, 2013. 11

[39] Maurizio Bocca, Anh Luong, Neal Patwari, and Thomas Schmid. Dial it in: Rotating RF sensors to enhance radio tomography. In *Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking*, pages 600–608, 2014. 11

[40] Dian Zhang and Lionel M Ni. Dynamic clustering for tracking multiple transceiver-free objects. In *IEEE International Conference on Pervasive Computing and Communications, 2009*, pages 1–8, 2009. 11

[41] Dian Zhang, Jian Ma, Quanbin Chen, and Lionel M Ni. An rf-based system for tracking transceiver-free objects. In *IEEE International Conference on Pervasive Computing and Communications,*, pages 135–144, 2007. 11

[42] Dian Zhang, Yunhuai Liu, Xiaonan Guo, and Lionel M Ni. Rass: A real-time, accurate, and scalable system for tracking transceiver-free objects.

IEEE Transactions on Parallel and Distributed Systems, 24(5):996–1008, 2013. 11

[43] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. Lifs: low human-effort, device-free localization with fine-grained subcarrier information. In *Proceedings of the 22nd ACM Annual International Conference on Mobile Computing and Networking*, pages 243–256, 2016. 12, 147

[44] Xiang Li, Shengjie Li, Daqing Zhang, Jie Xiong, Yasha Wang, and Hong Mei. Dynamic-music: accurate device-free indoor localization. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 196–207, 2016. 12, 147

[45] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *IEEE transactions on Antennas and Propagation*, 34(3):276–280, 1986. 12

[46] Kiran Joshi, Dinesh Bharadia, Manikanta Kotaru, and Sachin Katti. Wideo: fine-grained device-free motion tracing using rf backscatter. In *Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation*, pages 189–204, 2015. 12

[47] Feng Hong, Xiang Wang, Yanni Yang, Yuan Zong, Yuliang Zhang, and Zhongwen Guo. WFID: passive device-free human identification using WiFi signal. In *Proceedings of the 13th ACM International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pages 47–56, 2016. 14

[48] Tong Xin, Bin Guo, Zhu Wang, Mingyang Li, Zhiwen Yu, and Xingshe Zhou. Freesense: Indoor human identification with Wi-Fi signals. In *IEEE Global Communications Conference*, pages 1–7, 2016. 14

[49] Xinguo Yu. Approaches and principles of fall detection for elderly and patient. In *10th IEEE International Conference on e-health Networking, Applications and Services, 2008.*, pages 42–47, 2008. 15

[50] Ge Wu. Distinguishing fall activities from normal activities by velocity characteristics. *Journal of biomechanics*, 33(11):1497–1500, 2000. 15

[51] Thomas Degen, Heinz Jaeckel, Michael Rufer, and Stefan Wyss. Speedy: A fall detector in a wrist watch. In *Proceedings of the IEEE International Symposium on Wearable Computers*, page 184. IEEE Computer Society, 2003. 15

[52] Alan K. Bourke, Karol J O'Donovan, John Nelson, and Gearoid M O'Laighin. Fall-detection through vertical velocity thresholding using a tri-axial accelerometer characterized using an optical motion-capture system. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2832–2835, 2008. 15

[53] Stefano Abbate, Marco Avvenuti, Guglielmo Cola, Paolo Corsini, Janet Light, and Alessio Vecchio. Recognition of false alarms in fall detection systems. In *IEEE Consumer Communications and Networking Conference*, pages 23–28, 2011. 15

[54] Alan K. Bourke, Karol J O'Donovan, and Gearoid M O'laighin. The identification of vertical velocity profiles using an inertial sensor to investigate pre-impact detection of falls. *Medical engineering & physics*, 30(7):937–946, 2008. 15

[55] Ge Wu and Shuwan Xue. Portable preimpact fall detector with inertial sensors. *IEEE Transactions on neural systems and Rehabilitation Engineering*, 16(2):178–183, 2008. 15

[56] Federico Bianchi, Stephen J Redmond, Michael R Narayanan, Sergio Cerutti, and Nigel H Lovell. Barometric pressure and triaxial accelerometry-based falls event detection. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6):619–627, 2010. 15

[57] A Moncada-Torres, K Leuenberger, R Gonzenbach, A Luft, and R Gassert. Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiological measurement*, 35(7):1245, 2014. 15

[58] Stefano Abbate, Marco Avvenuti, Francesco Bonatesta, Guglielmo Cola, Paolo Corsini, and Alessio Vecchio. A smartphone-based fall detection system. *Pervasive and Mobile Computing*, 8(6):883–899, 2012. 15

[59] B. Aguiar, T. Rocha, J. Silva, and I. Sousa. Accelerometer-based fall detection for smartphones. In *IEEE International Symposium on Medical Measurements and Applications*, pages 1–6, June 2014. 15

[60] L. J. Kau and C. S. Chen. A smart phone-based pocket fall accident detection, positioning, and rescue system. *IEEE Journal of Biomedical and Health Informatics*, 19(1):44–56, Jan 2015. 15

[61] Apple Watch 4 Is Now An FDA Class 2 Medical Device: Detects Falls, Irregular Heart Rhythm. `https://www.forbes.com/sites/jeanbaptiste/2018/09/14/apple-watch-4-is-now-an-fda-class-2-medical-device-detects-falls-irregular-heart-rhythm/`, 2018. Accessed: 2018-09-17. 15

[62] Medical Alert Advice. Automatic fall detection for seniors. `https://www.medicalalertadvice.com/fall-detection/`, 2018. Accessed: 2018-01-23. 15

[63] Alan K. Bourke et al. Fall detection algorithms for real-world falls harvested from lumbar sensors in the elderly population: A machine learning approach. In *Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pages 3712–3715, 2016. 15, 139

[64] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Transactions on Information Technology in Biomedicine*, 10(1):156–167, 2006. 15, 16, 126

[65] T. Shany, S. J. Redmond, M. R. Narayanan, and N. H. Lovell. Sensors-based wearable systems for monitoring of human movement and falls. *IEEE Sensors Journal*, 12(3):658–670, March 2012. 15

[66] N Noury, P Barralon, G Virone, P Boissy, M Hamel, and P Rumeau. A smart sensor based on rules and its evaluation in daily routines. In *Proceedings of the 25th annual international conference of the IEEE Engineering in medicine and biology society*, volume 4, pages 3286–3289, 2003. 16

[67] Tony Rosen, Karin A Mack, and Rita K Noonan. Slipping and tripping: fall injuries in adults associated with rugs and carpets. *Journal of injury and violence research*, 5(1):61, 2013. 16

[68] Jane Fleming and Carol Brayne. Inability to get up after falling, subsequent time on floor, and summoning help: prospective cohort study in people over 90. 337, 2008. 16

[69] Brad Mager, Neal Patwari, and Maurizio Bocca. Fall detection using rf sensor networks. In *IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, pages 3472–3476, 2013. 16, 20, 21

[70] L Day. Falls in older people: Risk factors and strategies for prevention. *Injury Prevention*, 9(1):93, 2003. 16

[71] Anh Luong, Alemayehu Solomon Abrar, Thomas Schmid, and Neal Patwari. RSS step size: 1 dB is not enough! In *Procdings of the 3rd ACM Workshop on Hot Topics in Wireless*, pages 17–21, 2016. 17, 20

[72] Paul Edward Cuddihy, Jeffrey Michael Ashe, Corey Nicholas Bufi, and Sahika Genc. Radar based systems and methods for detecting a fallen person, 2014. US Patent 8,742,935. 17

[73] Leon Cohen. *Time-frequency Analysis: Theory and Applications*. Prentice-Hall, Inc., 1995. ISBN 0-13-594532-1. 18

[74] Liang Liu, Mihail Popescu, Marjorie Skubic, Marilyn Rantz, Tarik Yardibi, and Paul Cuddihy. Automatic fall detection based on doppler radar motion signature. In *5th IEEE International Conference on Pervasive Computing Technologies for Healthcare*, pages 222–225, 2011. 18

[75] Branka Jokanovic, Moeness G Amin, and Fauzia Ahmad. Effect of data representations on deep learning in fall detection. In *IEEE Sensor Array and Multichannel Signal Processing Workshop*, pages 1–5, 2016. 18

[76] Boualem Boashash, Nabeel Ali Khan, and Taoufik Ben-Jabeur. Time–frequency features for pattern recognition using high-resolution tfds: A tutorial review. *Digital Signal Processing*, 40:1–30, 2015. 18, 124, 125

[77] Baris Erol, Moeness Amin, Fauzia Ahmad, and Boualem Boashash. Radar fall detectors: A comparison. In *Radar Sensor Technology*, volume 9829, page 982918. International Society for Optics and Photonics, 2016. 18, 76, 79, 82, 121

[78] Qisong Wu et al. Radar-based fall detection based on Doppler time–frequency signatures for assisted living. *IET Radar, Sonar & Navigation*, 9(2):164–172, 2015. 18, 78

[79] Ph Van Dorp and FCA Groen. Feature-based human motion parameter estimation with radar. *IET Radar, Sonar & Navigation*, 2(2):135–145, 2008. 18, 19, 78, 82, 103

[80] Baris Erol, Moeness Amin, Zhichong Zhou, and Jun Zhang. Range information for reducing fall false alarms in assisted living. In *IEEE Radar Conference*, pages 1–6, 2016. 19

[81] Haofei Wang, Linyun Ren, Erke Mao, and Aly E Fathy. Phase based motion characteristics measurement for fall detection by using stepped-frequency continuous wave radar. In *IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems*, pages 43–45, 2016. 19

[82] Tong Liu, Jun Liu, and Xiao-mu Luo. Radio tomographic imaging based body pose sensing for fall detection. *Journal of Ambient Intelligence and Humanized Computing*, 5(6):897–907, 2014. 20

[83] Sanaz Kianoush, Stefano Savazzi, Federico Vicentini, Vittorio Rampa, and Matteo Giussani. Device-free rf human body fall detection and localization

in industrial workplaces. *IEEE Internet of Things Journal*, 4(2):351–362, 2017. 20

[84] Zicheng Chi, Y. Yao, Tiantian Xie, Zhichuan Huang, M. Hammond, and Ting Zhu. Harmony: Exploiting coarse-grained received signal strength from IoT devices for human activity recognition. In *IEEE 24th Int. Conf. on Network Protocols*, pages 1–10, 2016. 20

[85] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. Contactless Respiration Monitoring Via Off-the-Shelf WiFi Devices. *IEEE Transactions on Mobile Computing*, 15(10):2466–2479, 2016. 21

[86] Yan Wang et al. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *Procedings of the 20th ACM Annual Int. Conf. on Mobile Computing and Networking*, pages 617–628, 2014. 21, 114

[87] F. Xiao et al. Seare: A system for exercise activity recognition and quality evaluation based on green sensing. *IEEE Transactions on Emerging Topics in Computing*, 2018. 21

[88] Guanhua Wang, Yongpan Zou, Zimu Zhou, Kaishun Wu, and Lionel M. Ni. We Can Hear You with Wi-Fi! In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, 2014. 21

[89] Li Sun, Souvik Sen, Dimitrios Koutsonikolas, and Kyu-Han Kim. Widraw: Enabling hands-free drawing in the air on commodity wifi devices. In *Proceedings of the 21st ACM Annual International Conference on Mobile Computing and Networking*, pages 77–89, 2015. 21

[90] Kamran Ali, Alex X. Liu, Wei Wang, and Muhammad Shahzad. Keystroke Recognition Using WiFi Signals. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 90–102, 2015. 21

[91] Sheng Tan and Jie Yang. WiFinger: leveraging commodity WiFi for fine-grained finger gesture recognition. In *Proceedings of the 17th ACM Inter-*

national Symposium on Mobile Ad Hoc Networking and Computing, pages 201–210, 2016. 21

[92] Nan Yu, Wei Wang, Alex X. Liu, and Lingtao Kong. QGesture: Quantifying Gesture Distance and Direction with WiFi Signals. *Procedings ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(1):51:1–51:23, 2018. 21

[93] Yongsen Ma, Gang Zhou, Shuangquan Wang, Hongyang Zhao, and Woosub Jung. SignFi: Sign Language Recognition Using WiFi. *Procedings of ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(1):23:1–23:21, 2018. 21

[94] Jiacheng Shang and Jie Wu. A Robust Sign Language Recognition System with Multiple Wi-Fi Devices. In *Proceedings of the Workshop on Mobility in the Evolving Internet Architecture*, pages 19–24, 2017. 21

[95] Kun Qian et al. Inferring Motion Direction using Commodity Wi-Fi for Interactive Exergames. In *Procedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017. 21, 22, 147

[96] Daqing Zhang, Hao Wang, and Dan Wu. Toward centimeter-scale human activity sensing with Wi-Fi signals. *Computer*, 50(1):48–57, 2017. 21, 134

[97] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. Human Respiration Detection with Commodity Wifi Devices: Do User Location and Body Orientation Matter? In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016. 21

[98] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. Wi-Sleep: Contactless sleep monitoring via WiFi signals. In *IEEE Real-Time Systems Symposium*, pages 346–355, 2014. 21

[99] Yuxi Wang et al. Wifall: Device-free fall detection by wireless networks. *IEEE Transactions on Mobile Computing*, 16(2):581–594, 2017. 21, 23, 114

[100] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li. RT-Fall: A Real-Time and Contactless Fall Detection System with Commodity WiFi

Devices. *IEEE Transactions on Mobile Computing*, 16(2):511–526, 2017. 21, 23, 79, 84, 93, 103, 114, 138

[101] Wei Wang et al. Gait Recognition Using WiFi Signals. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016. 21, 22, 69, 74, 78, 147

[102] Kun Qian et al. Widar: decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi. In *Procedings of the 18th ACM Int. Symp. on Mobile Ad Hoc Networking and Computing*, page 6, 2017. 21, 69, 118, 124

[103] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM Computer Communication Review*, pages 269–282, 2015. 29

[104] Souvik Sen, Romit Roy Choudhury, Bozidar Radunovic, and Tom Minka. Precise indoor localization using phy layer information. In *Proceedings of the 10th ACM Workshop on hot topics in networks*, page 18, 2011. 30

[105] Yaxiong Xie, Zhenjiang Li, and Mo Li. Precise power delay profiling with commodity wifi. In *Proceedings of the 21st ACM Annual International Conference on Mobile Computing and Networking*, pages 53–64, 2015. 30

[106] Yiwei Zhuo, Hongzi Zhu, and Hua Xue. Identifying a New Non-Linear CSI Phase Measurement Error with Commodity WiFi Devices. In *IEEE 22nd International Conference on Parallel and Distributed Systems*, pages 72–79, 2016. 30

[107] H. Hashemi. The indoor radio propagation channel. *Procedings of the IEEE*, 81(7):943–968, 1993. 31

[108] Kaishun Wu et al. Fila: Fine-grained indoor localization. In *Procedings of IEEE Conference on Computer Communications*, 2012. 31

[109] Deepak Vasisht et al. Decimeter-level localization with a single WiFi access point. In *Procedings of the NSDI*, 2016. 32

[110] Souvik Sen et al. Precise indoor localization using phy layer information. In *Procedings of the 10th ACM Workshop on Hot Topics in Networks*, 2011. 32

[111] Neal Patwari and Joey Wilson. People-Sensing Spatial Characteristics of RF Sensor Networks. *CoRR*, abs/0911.1972, 2009. 32

[112] N. Patwari and J. Wilson. RF Sensor Networks for Device-Free Localization: Measurements, Models, and Algorithms. *Procedings of the IEEE*, 98(11): 1961–1973, 2010. 34, 37

[113] Daniel Halperin et al. Predictable 802.11 packet delivery from wireless channel measurements. *ACM SIGCOMM Comput. Commun. Rev.*, 41(4), August 2010. 34

[114] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In *International Conference on Artificial Neural Networks*, pages 583–588. Springer, 1997. 49

[115] Sebastian Mika et al. Kernel PCA and De-noising in Feature Spaces. In *Procedings of Advances in Neural Information Processing Systems II*, 1999. 49

[116] Quan Wang. Kernel Principal Component Analysis and its Applications in Face Recognition and Active Shape Models. *CoRR*, abs/1207.3538, 2012. 49

[117] Alex J Smola et al. The connection between regularization operators and support vector kernels. *Neural networks*, 11(4):637–649, 1998. 50

[118] David J Ketchen and Christopher L Shook. The application of cluster analysis in strategic management research: an analysis and critique. *Strategic management journal*, 17(6):441–458, 1996. 53

[119] Daniel Halperin et al. Tool release: gathering 802.11n traces with channel state information. *ACM SIGCOMM Comput. Commun. Rev.*, 41(1):53–53, 2011. 56, 87

[120] Sameera Palipana, David Rojas, Piyush Agrawal, and Dirk Pesch. FallDeFi: Ubiquitous Fall Detection Using Commodity Wi-Fi Devices. *Procdings ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4):155:1–155:25, January 2018. 69, 124, 129

[121] Bo Yu Su, KC Ho, Marilyn J Rantz, and Marjorie Skubic. Doppler radar fall activity detection using the wavelet transform. *IEEE Transactions on Biomedical Engineering*, 62(3):865–875, 2015. 71, 79, 84

[122] Youngwook Kim and Hao Ling. Human activity classification based on micro-doppler signatures using a support vector machine. *IEEE Transactions on Geoscience and Remote Sensing*, 47(5):1328–1337, 2009. 78

[123] Luis Ramirez Rivera et al. Radar-based fall detection exploiting time-frequency features. In *IEEE China Summit & International Conference on Signal and Information Processing*, pages 713–717, 2014. 79, 82

[124] Baris Erol, Moeness G Amin, Boualem Boashash, Fauzia Ahmad, and Yimin D Zhang. Wideband radar based fall motion detection for a generic elderly. In *IEEE Asilomar conference on Signals, Systems and Computers*, pages 1768–1772, 2016. 80

[125] Aihua Zhang, Bin Yang, and Ling Huang. Feature Extraction of EEG Signals Using Power Spectral Entropy. In *Procdings of the 2008 Int. Conf. on BioMedical Engineering and Informatics - Volume 02*, 2008. 83

[126] Dierk Schleicher. Hausdorff dimension, its properties, and its surprises. *The American Mathematical Monthly*, 114(6):509–528, 2007. 83

[127] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology*, 2(3): 27, 2011. 85

[128] Sameera Palipana. Falldefi source code and data. `https://github.com/dmsp123/FallDeFi`, 2017. 87

[129] David L Donoho. De-noising by soft-thresholding. *IEEE transactions on information theory*, 41(3):613–627, 1995. 90

[130] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002. 93, 126

[131] Yonglong Tian, Guang-He Lee, Hao He, Chen-Yu Hsu, and Dina Katabi. Rf-based fall monitoring using convolutional neural networks. *Proceedings ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(3):137:1–137:24, 2018. 114

[132] Songwon Seo. *A review and comparison of methods for detecting outliers in univariate data sets*. PhD thesis, University of Pittsburgh, 2006. 119

[133] William S Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368):829–836, 1979. 122

[134] Baris Erol et al. Wideband radar based fall motion detection for a generic elderly. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 1768–1772, 2016. 124

[135] Baris Erol and Moeness G Amin. Fall motion detection using combined range and doppler features. In *Signal Processing Conference, 2016 24th European*, pages 2075–2080, 2016. 147

[136] Manuel Montero-Odasso, Marcelo Schapira, Enrique R Soriano, Miguel Varela, Roberto Kaplan, Luis A Camera, and L Marcelo Mayorga. Gait velocity as a single predictor of adverse events in healthy seniors aged 75 years and older. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60(10):1304–1309, 2005. 147

[137] Gabor Abellan Van Kan, Y Rolland, S Andrieu, J Bauer, O Beauchet, M Bonnefoy, M Cesari, LM Donini, S Gillette-Guyonnet, M Inzitari, et al. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an international academy on nutrition and aging (iana) task force. *The journal of Nutrition, Health & Aging*, 13(10):881–889, 2009. 147

[138] Stacy Fritz and Michelle Lusardi. White paper:walking speed: the sixth vital sign. *Journal of Geriatric Physical Therapy*, 32(2):2–5, 2009. 147

[139] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):72, 2017. 147

[140] Bo Wei, Wen Hu, Mingrui Yang, and Chun Tung Chou. Radio-based device-free activity recognition with radio frequency interference. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 2015. 148