LEVERAGING CHANNEL STATE INFORMATION FROM COTS WIFI
ROUTER TO DETECT WATER FLOW PATTERN

by

Abhinav Kumar

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Technology

Charlotte

2019

Approved by:

_____
Dr. Weichao Wang

_____
Dr. Pu Wang

_____
Dr. Yu Wang

ABSTRACT

ABHINAV KUMAR. Leveraging channel state information from cots Wi-Fi router to detect water flow pattern. (Under the direction of DR. WEICHAO WANG)

Wi-Fi based activity recognition serves a multitude of applications in fields such as smart homes, health care, and security. The property of the channel state information to provide fine-grained information has been utilized in the above-mentioned fields. However, we wanted to diversify the use of channel state information to even broader areas and that's why used it to detect the water flow and differentiate between different water patterns. Given the huge scarcity of water in the current world our work tends to serve a practical purpose. The basic idea is to utilize the signal variation caused due to the water flow pattern. The static objects such as furniture, still human causes reflection of signal whereas dynamic or moving objects causes additional propagation paths. These additional propagation paths can be observed by measuring the channel state information amplitude between the two routers.

There have been a number of challenges on the way which needed to be figured out before we can get a satisfactory result. Some of them include removing the background noise from the CSI data collected, selecting a classifier which can accurately differentiate between the different patterns of the water flow. We performed various signal processing techniques to reduce the noise and to get a much better representation of the CSI waveform. The Multi-class SVM classifier was modeled and trained to predict the accuracy of different labels collected. Our model achieved 90.35 % accuracy in classifying different labels. Considering this as base work for the detection of water flow pattern, accuracy can be improved later on with some additional functionalities.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CSI   Channel State Information

COTS   Commercial Off-The-Shelf

RSSI   Received Signal Strength Indicator

NIC   Network Interface Card

LOS   Line-of-Sight

USRP   Universal Software Radio Peripheral

ADC   Analog-to-Digital Converter

OFDM   Orthogonal Frequency-Division Multiplexing

MCS   Modulation and Coding Scheme

PCA   Principal Component Analysis

MIMO   Multiple-Input Multiple-Output

FFT   Fast Fourier Transform

CFO   Carrier Frequency Offset

SVM   Support Vector Machine

RBF   Radial Basis Function

CHAPTER 1: INTRODUCTION

The CSI based activity recognition has been a hot topic for research in recent years. There have been many projects done leveraging the physical layer channel state information from the NICs (network interface cards). Some of the prominent ones include human activity recognition [1], gait recognition [2], and keystroke recognition [3]. We proposed a novel approach here to use the existing CSI based technology for detection of water flow pattern. This technology can be leveraged by firefighters to detect the effectiveness of their operation in extreme conditions. Farmers can utilize it for irrigation purposes. For evaluation of the model, we used multi-class SVM classifier which can differentiate between various settings of the water flow from a showerhead.

1.1 Leveraging Wi-Fi signal for water flow recognition

In the modern world, one of the major problems faced by many is the scarcity of water. The natural resources of water are getting dried up and the exponential rise in population is only making the matter worse. There have been various efforts made to preserve this vital natural resource but still, a lot needs to be done. Ours is an innovative approach designed to detect the water flow pattern with the help of CSI data. It can be used to help firefighters in their effort also farmers can use it for the irrigation purposes.

Currently, the firefighters risk their lives during an operation. Using our approach, they can monitor the different zones in which the water is sprayed and increase the efficiency of their operation. It can play a vital role in saving their lives. Farmers can use our technology to better monitor the different regions in which the water is sprayed. It can help them to identify the regions which did not get sufficient amount of water.

Recently there has been an increase in the research work related to wireless sensing. It has been used to detect human activity, fall detection, hand gesture and many other forms of detection. Here we tried to leverage the wireless sensing technology to detect water flow pattern by measuring the variation in CSI amplitude. Video surveillance is another method to detect water flow used in many places. But it has privacy concerns associated with it. Keeping all these things in minds the wireless surveillance is much more efficient and doesn't have any privacy concerns.

In Figure 1, we show a basic framework as to how we are going to conduct our experiment. There are two routers one acting as an access point while the other one as a client and in between a shower head is used to flow the water in LOS of the router.

Wi-Fi router
Access Point

Water flowing from
a shower head

Wi-Fi router
Client

Figure 1 Basic framework of our experiment

1.2 Motivation behind the use of Wi-Fi data

As evident there is not much option to detect the flow of water, video surveillance can be used but it covers many aspects and is not feasible for detection of water flow. It also requires significant installation cost, has privacy concerns which may cause annoyance to many users and its performance decreases drastically under bad lighting conditions in

spite of having a significant image recognition algorithm available. Overall the use of Wi-Fi is much more efficient, cost-effective and requires much less maintenance.

As we can see there has not been much research done in the detection of water flow and the existing method has many drawbacks as mentioned in the above scenario. So, to overcome those we are using Wi-Fi data which is easily available from the wireless routers. There are a lot of advantages of using it. Some of the prominent ones include, it is easily available nowadays whether it is home or office. It can pass through walls and so can even work in non-line of sight scenario. It preserves the privacy of the user. Unlike video surveillance system it can perform effectively even at night under dark conditions. It does not require users to wear any devices such as Fitbit or smartwatch. As many times user forget to wear them and then it can't measure any activity.

1.3 Related work using wireless data

Due to its wide application potential and rapid development in wireless technologies, it has been used in a number of scenarios. There has been a lot of research in recent times using the wireless signal for activity recognition. It can be mainly classified in these categories:

1.3.1 RSSI Based

RSSI (Received Signal Strength Indicator) is used to measure the strength of the signal from a connected wireless device. It detects the signal quality from a connected client device. It measures the relative strength and that's why has no absolute value. The RSSI value differs between the chipset manufacturer. Such as, Atheros uses a scale of 0-60 whereas Cisco uses a scale of 0-100. RSSI based activity recognition systems leverage the received signal strength changes caused by human activities [4]. This approach is not

very effective since the received signal strength from the commercial Wi-Fi devices is not very strong and so only provides coarse-grained information about the activities. Still, it has been used in a number of scenarios including localization of device [5] also it has been used in recognizing activity with low accuracy like here the accuracy is 56% for 7 different gestures [6].

1.3.2 Specialized Hardware Based

Specially designed hardware such as USRP radios provides fine-grained wireless information about human activity. WiTrack [7] uses USRP radios to track the 3D motion of a user using radio reflection that bounces off his body. It can also detect the direction of hand motion; it uses one antenna for transmission and three antennas for the receiver. AllSee using a specially designed analog circuit to extract the amplitude information of received signals has been shown to recognize gesture within a short distance of 2.5 feet [8].

1.3.3 CSI Based

Recently CSI has been used in human activity recognition [1], indoor localization [9], gait recognition [2] and various other work. The high number of research using CSI data is being attributed to its ease of availability from various commercial devices such as Intel 5300 and Atheros Network Interface Cards(NICs). CSI provides fine-grained information of the wireless signal and hence is much more accurate than the RSSI based detection. Han et al. proposed to use CSI data for fall detection, which can be utilized for the elderly [10]. Xi et al. proposed to use CSI data to count the number of people in a crowd [11]. DeMan [12] built a unified framework with commodity Wi-Fi devices to detect both stationary and moving persons using the maximum eigenvalues.

Here, we are using the variation in CSI data to detect the flow of water and different water patterns. Noisy CSI data is being filtered using PCA based technique. PCA captures the most representative signal of water flow pattern; after which we used a multi-class SVM classifier to differentiate between the various sets of the water flow patterns.

1.4 Technical Challenges in CSI

There are a number of technical challenges in using CSI data for activity recognition. The CSI data collected on the receiver side is very noisy and contains a lot of unwanted information and therefore is not suitable for any kind of computation.

The second challenge includes the selection of features, it is not easy to decide which feature to select as the feature selected should be unique and able to differentiate between similar looking CSI amplitude.

The third major challenge associated with the use of CSI includes building a classifier which can accurately differentiate between different scenarios and give an appropriate result. Especially when there are many classifiers which can provide similar results.

1.5 Thesis organization

➤ Chapter 2 covers the literature review and introduction to CSI. We go through some of the recent works in wireless signal processing utilizing the CSI data. The foundation of CSI is covered, and question about how CSI is able to collect such fine-grained information for activity recognition is answered. We discuss MIMO technology and see how it has revolutionized the use of wireless system.

- Chapter 3 mainly deals with the CSI data pre-processing stage. The raw CSI data contains a lot of noise which needs to be filtered out before it can be used for any further processing. We introduced the Hampel identifier to remove the outliers from the CSI data. After that, we used different filters to remove the noise from the CSI data. At last PCA technique is used for dimensionality reduction as it provides optimum results.

- Chapter 4 covers the feature extraction from the filtered CSI data. After that, the process of building a multi-class SVM classifier was mentioned. Python scikit-learn library was used to build and train the classifier. We also discussed the method for dataset collection and how the experiment was set-up and performed.

- Chapter 5 is the last chapter and mainly summarizes our work and the result obtained. We also discussed some of the future work we plan to implement based on the current model.

CHAPTER 2: LITERATURE REVIEW AND CSI INTRODUCTION

This chapter will provide a comprehensive literature review on the work done in the wireless system for various detection method and indoor based localization. The first part of the literature review will go through a brief introduction to the wireless communication system and will discuss some key terms and technologies of the wireless communication system. In the second part, we will introduce channel state information (CSI), its uses, and the tool used for the collection of CSI data. We will also cover multiple-input multiple-output (MIMO) system which results in a more robust wireless signal-based detection method as compared to previously used received signal strength (RSS) technique. As discussed in the previous chapter the wireless detection system provides much more perks as compared to other forms of detection techniques. It has led to an increase in the number of research work in wireless technologies.

After that, we cover the architecture of our design which enables us to observe the variations due to flowing water from the shower head. Here only the CSI amplitude is used for all kind of detection purposes. Most of the earlier research on CSI data uses the Intel 5300 chip, with the tool developed by Dan. Halperin [13] to get the CSI data with 30 subcarriers. However, here we are using an Atheros CSI tool [14] developed by Xie et al., it is relatively new and provides more subcarriers for both 20 MHz and 40 MHz bandwidth. The tool can be installed on NIC having Atheros chip or a few particular Qualcomm chip. Here we installed the tool on top of two TP-Link routers.

The CSI signal collected at the receiver side contains a lot of background noises and had to be pre-processed before it can be used for any relevant tasks. Traditional method includes low pass-filter to attenuate the noise but those are not very effective.

That's why here Wavelet-based de-noising technique is used. It preserves the most relevant CSI variation caused due to water flow pattern while removing the unwanted noise [15].

2.1 MIMO system

The multiple-input multiple-output system is a radio communication technology that is being used in many new technologies these days. MIMO system have multiple antennas elements at both the transmitter and the receiver side. They were first investigated through computer simulation during the 1980s [16]. Since then the interest in the MIMO system has exploded and it has become much more diversified to be used in various equipment. There is a large number of MIMO routers available in the market and its demand has only been on the rise.

The increase in the number of antenna elements allows for higher data rates, a better quality of service and higher network capacity [17]. As compared to wired technology it offers easier deployment, lower costs and greater flexibility. The introduction of MIMO has brought a new dimension called 'space' to digital communication systems, in addition to the 'time' dimension which is considered as a natural dimension in the digital communication systems. That's why MIMO technology is also known as 'space-time' wireless technology [18]. Figure 2 shows how spatial multiplexing is used to increase the data traffic by utilizing different paths between a transmitter and a receiver.
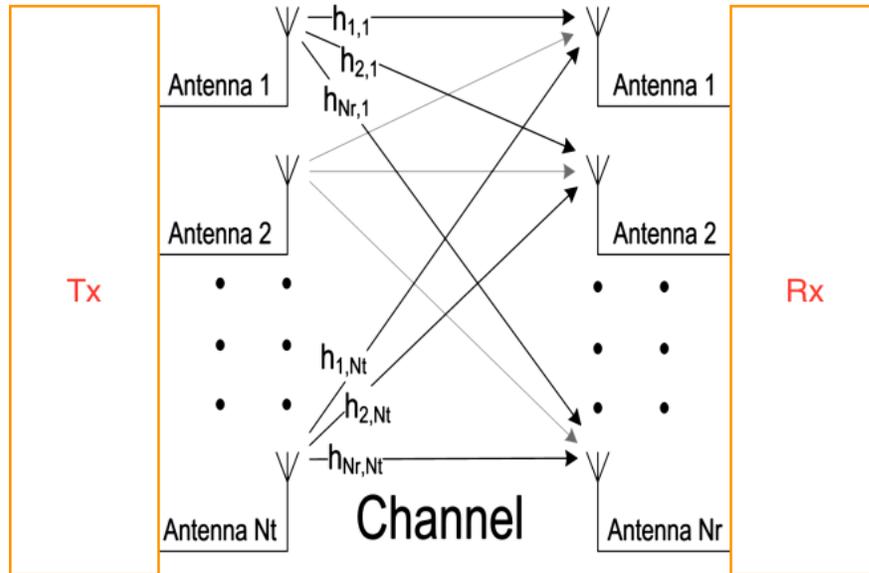
Figure 2: Example of a MIMO system

In spite of all the benefits of a MIMO system, there are some disadvantages. It increases the hardware cost and radio frequency circuit consumption. The component which is most power-hungry is the high-resolution analog-to-digital converter (ADC) since its power consumption is scaled exponentially with the number of quantization bits and linearly with the baseband bandwidth [19]. Low-resolution ADCs is used to overcome this challenge and is quite popular due to its lower hardware complexity.

One unique feature of MIMO system is spatial multiplexing which is used to increase the data rate by transmitting different streams of information in a parallel channel. The number of spatial multiplexing channels depend on the number of antennas, given by the following equation:

$$Ns = \min (Nt, Nr) \tag{2.1}$$

Here, Nt represents the number of transmitting antennas and Nr represents the number of receiving antennas.

Modulation and coding schemes (MCS) are used to determine the throughput rate of an orthogonal frequency division multiplexing (HT-OFDM) starting from IEEE 802.11n. HT-OFDM uses various parameters such as channel size, coding method, number of spatial streams, modulation technique and guard interval. Currently, 802.11n have 77 different types of MCS index for both 20 and 40 MHz channels. Here, MCS index of 15 is being used that has two spatial streams.

2.2 Channel State Information (CSI)

In wireless system channel state information (CSI) refers to known channel properties of a communication link. It provides the state of a communication link from the transmitting side to the receiving side. It has been used in various activities recognition and captures scattering, fading and power decay of signal with distance. When humans move, they cause scattering of Wi-Fi signals which can be revealed through the properties of channel state information. CSI can be exported from commercially available wireless Network Interface Card (NIC) and is therefore used in a number of researches nowadays [20]. CSI is much more stable and provides more information as compared to RSSI because CSI is measured per orthogonal frequency-division multiplexing (OFDM) from each packet. Whereas, RSSI is measured by a single value per packet. CSI is also more resilient toward a change in the environment [21].

To collect the CSI data using two routers, the firmware of the router here, TP-Link-TL WR1043ND is upgraded to OpenWRT by replacing the custom firmware using a serial port connection. Atheros CSI tool is being installed on top of the OpenWRT firmware which allows to fetch the valuable CSI data from the router. Mathematically, CSI can be modeled in the frequency domain as

$$y = \mathbf{H}.x + n \tag{2.2}$$

Where y is the received signal, x the transmitted signal, H denotes the complex channel matrix consisting of CSI values and n represents the noise present in the path.

The OFDM technique processes the CSI value of 56 subcarriers at 20 MHz bandwidth and 114 subcarriers at 40 MHz bandwidth. Our platform uses 20 MHz bandwidth when both the transmitter and receiver routers are operating at 2.4 GHz frequency. Each subcarrier provides the amplitude and phase information as a complex number about the path taken, represented by the below equation.

$$H_k = |H_k| \, e^{(j\angle H_k)} \tag{2.3}$$

Where $|H_k|$ and $\angle H_k$ represents the amplitude and phase of CSI at the K_{th} subcarrier respectively.

2.2.1 OFDM technique

The CSI information comes from each subcarrier based on the OFDM technique. OFDM was first proposed by Chang [22]. In OFDM the overall spectrum band is subdivided into many small frequency bands known as subcarriers, illustrated in much more details in Figure 3. From the figure we can say that first on the transmission side, each carrier takes part of the data to be transmitted on an OFDM carrier signal. Data is performed using Inverse Fast Fourier Transform (IFFT) to transmit through the air. IFFT on data gives a set of complex time-domain samples that are further passed to the passband. The real and imaginary components are converted to an analog domain using digital-to-analog converters (DACs). The analog signals are then summed to give the transmission signal. On the receiver side, it samples them and pass them to a demodulation process chain and digitize them using analog to digital converters (ADCs).

The main difference between frequency division multiplexing(FDM) and OFDM is that OFDM system carriers are mutually packed and are orthogonal to other carriers. Orthogonal refers to that peak of one carrier occurs at the null of the other. That's why OFDM system is bandwidth efficient and provides higher data rates compared to FDM system in the same bandwidth usage.

One of the principal advantages of OFDM is its ability for transmission at near optimum in multipath channels. The main reason for that as is discussed in [23] is the insertion of small-time interval known as a guard interval which can eliminate intercarrier interference (ICI). The length of the guard interval is made equal to or greater than the time spread of the channel. But there are some downsides to using OFDM one of the most prominent one is sensitivity to frequency offset in the channel. Here, we are using MCS index of 15 with SGI (short guard interval) in 20 MHz bandwidth.
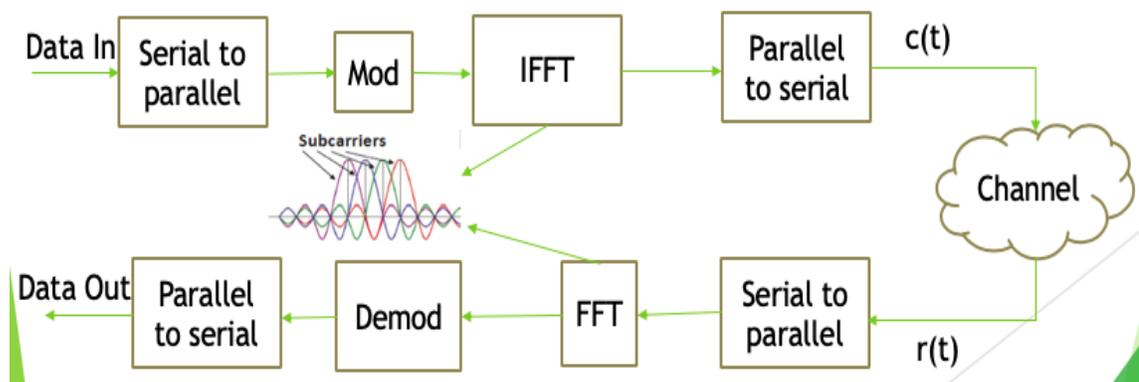
Figure 3 Working of CSI subcarriers using OFDM technique

2.2.2 Correlation between different OFDM subcarriers

In our experiment, we are utilizing 2.4GHz carrier frequency with a bandwidth of 20 MHz. Each CSI data packet contains 56 subcarriers which contain complex matrices with

dimension Ntx*Nrx. Here, Ntx and Nrx represent the number of transmitting and receiving antennas respectively. Here, we will see how the water flow pattern affects the amplitude of different subcarriers. Figure 4 depicts the amplitude of one subcarrier under different Tx-Rx pairs, i.e. subcarrier 1 in different TR link as can be seen from the diagram even the same subcarrier in different links show variation in shape and amplitude. However, as evident from Figure 5 CSI amplitude variation across different subcarriers in the same Tx-Rx pair is not much different. The subcarrier 1,2,3 and 4 displays very little variation in shape.
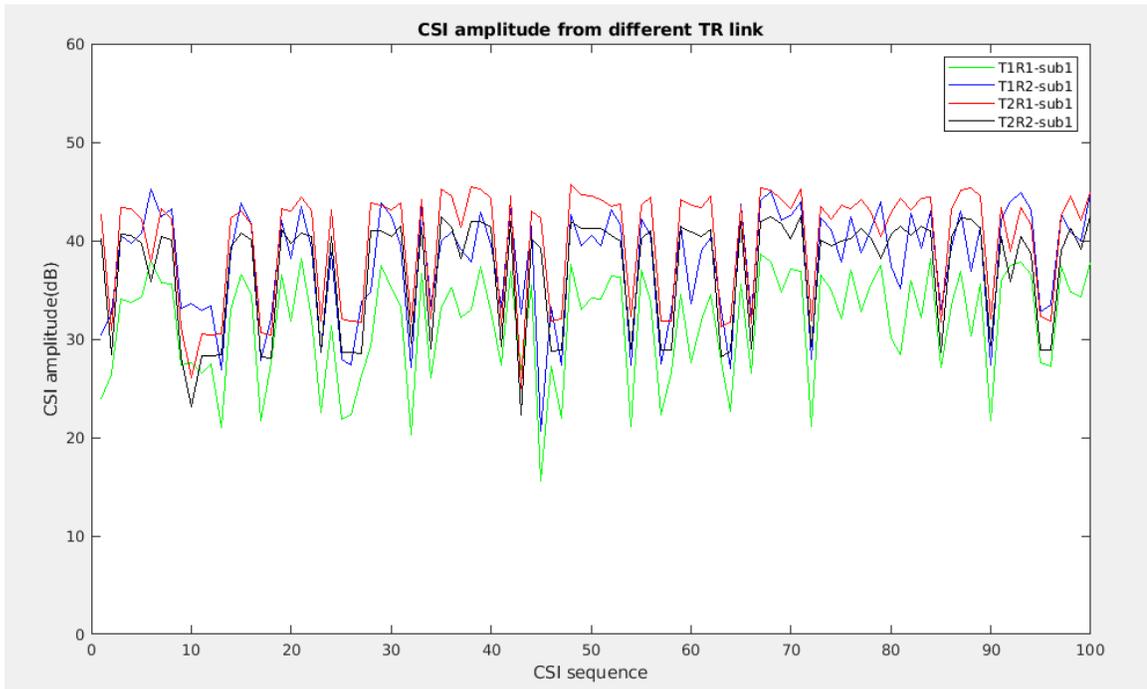
Figure 4 CSI amplitude for the same subcarrier (sub 1) against different TR links. As evident same subcarrier in different link show variation in pattern and amplitude.
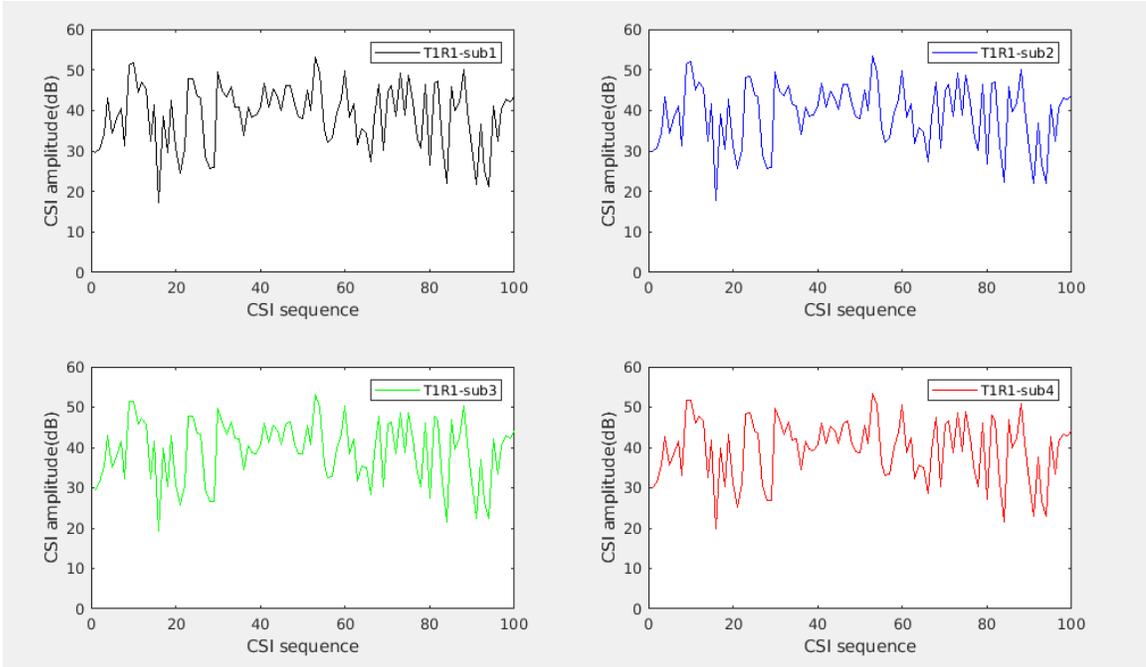
Figure 5 CSI amplitude of different subcarrier for the same TR link. As evident, it shows very little variance among different subcarriers.

2.3 Atheros CSI tool installation

Most of the research work uses two laptops equipped with Intel 5300 NIC, to extract the CSI data. It is a tool built to support extraction of CSI data and is done by modifying the Intel 5300 NIC with the permission of Intel. The Atheros CSI tool [14] is installed on top of an Atheros chip and not all devices support CSI extraction. Here, we installed the Atheros CSI tool on top of two TP-link routers by modifying the custom firmware to OpenWRT. Before doing that, we need to make sure about two things. First, the routers should be able to run OpenWRT distribution, and the Wi-Fi NIC of the devices should support CSI extraction. Here, we used TP-Link router (TL 1043ND) v3.0 which comes with chip 'QCA9558' and it supports CSI extraction.

The Atheros CSI tool provides much more detailed information about the path taken using more subcarriers and uses only two COTS Wi-Fi routers as compared to

14

conventional Intel 5300 NIC tool which requires a laptop equipped with Intel chip as a receiver. For example, in [24] for indoor based localization eleven laptops are used as receivers compared to that we are using just two COTS Wi-Fi routers equipped with the Atheros CSI tool.

During the upgrade of the router firmware, we encountered some problems such as while upgrading the old firmware image with the new OpenWRT image the router became unresponsive and it led to bricking of the router. To resolve it I had to create a serial connection using the serial port. For that, I had to open the router and then use a USB to TTL converter to connect the router serially. Finally, I was able to remove the faulty firmware and upgrade it with the correct one.

2.4 Summary

In this chapter, I went through the concepts which build the foundation for our tool. An introduction to wireless signal processing was presented in the first part. The proliferation of MIMO technology was discussed as to how it increases the data rate by using multiple antennas on both sender and receiver side. The concepts of CSI and its uses were discussed, and how CSI uses OFDM technique to send multiple subcarriers orthogonal to each other were also covered. At last, we discussed the Atheros CSI tool used by us and some of the problems faced during the installation process.

CHAPTER 3: CSI DATA PRE-PROCESSING

This chapter will discuss various ways to pre-process the CSI data. The original CSI data contains lots of background noise and cannot be used for any useful information extraction. Here we will discuss how to resolve this first key challenge of extracting the useful information from the noisy CSI data. Before we can monitor any area of interest pre-processing of the data needs to be done by removing the noisy signal. The CSI phase data arriving at the receiver contains carrier frequency offset and is very unreliable, that's why here we ignore the CSI phase data and use only the CSI amplitude. In contrast to direct feature extraction from the CSI amplitude data of each subcarrier as done in [25] PCA (principal component analysis) based filtering technique proves much more efficient in extracting the more representative signals and that's why here PCA based technique is used.

3.1 Outlier removal based on Hampel Identifier

The cost-effective and easily available commercial Wi-Fi cards introduce various types of noises in the CSI measurement data. The internal state transitions produce noises caused due to transmission power, rate adaptation and thermal noises generated [26]. The carrier frequency offset is introduced due to the difference in the oscillator of transmitter/receiver hardware. It leads to random variation in the phase of CSI [14]. To address the issue of CFO (carrier frequency offset) we ignore the CSI phase data. The CSI amplitude data is also affected by surrounding electromagnetic noises.

To get rid of the variation or outliers which are not caused by flowing water we use Hampel Identifier algorithm [27]. Generally, the sample mean $\overline{X}$ and variance S^2 of a

sample give a good estimation of data location. Here, the sample mean $\overline{X}$ gives an average of N data points and the variance S^2 summarizes the spread of data.

$$\overline{X} = \frac{\sum_{i=1}^{N} x_i}{N} , \qquad\qquad S^2 = \frac{\sum_{i=1}^{N}(x_i - \overline{X})^2}{n-1} \qquad\qquad (3.1)$$

But in the case when a dataset contains outliers, even a single out-of-scale observation can cause the sample mean to deviate significantly. Break-down point [28] was introduced by Hampel to measure the robustness of an estimator against outliers. Break-down point is the smallest percentage of outliers that can cause an estimator to take large aberrant values. Larger the breakdown points more robust it is. From the above equation, it can be seen that the sample mean has a breakdown point of 1/N. Thus, to estimate the location of an outlier, the median and the median absolute deviation (MAD) are often recommended.

Figure 6 below depicts a graph of CSI amplitude before and after the use of Hampel Identifier algorithm. In the figure, the outliers can be clearly seen in the CSI amplitude waveform highlighted with the square box. There are many outliers which are removed after the processing of data in Hampel Identifier algorithm. But still, it doesn't remove all the outliers. After that, we use the de-noising technique to get rid of the remaining outliers and de-noise the signal.
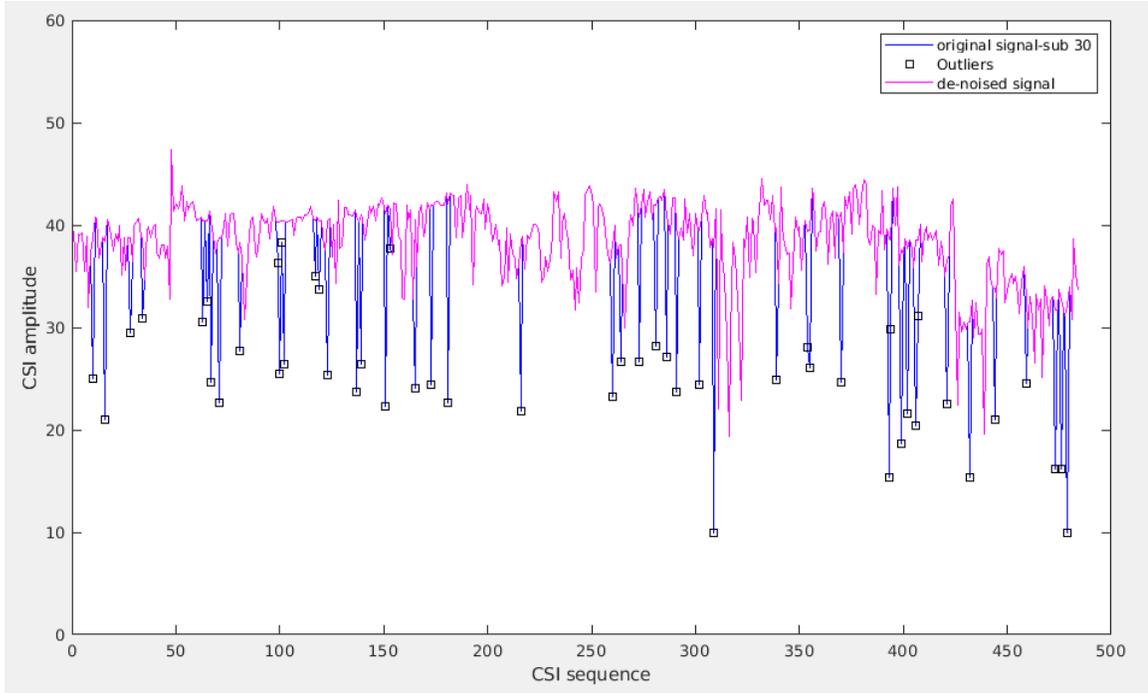
Figure 6 CSI sequence graph highlighting the outliers detected by Hampel identifier

For each value x of the base signal, the median of a window consisting of x and m/2 neighboring points on each side is computed. Then the standard deviation of x about its window is calculated. And if there is a difference by more than a predefined number of MAD, its value is replaced by the median. In our work for each value x of the signal, we varied the median window consisting of x and m neighboring points, i.e. m/2 on each side. So, we varied the value of m and observed where the maximum number of outliers are detected. For m = 10 we got the maximum outliers and that's why it was selected.

3.2 Noise-filtering: Wavelet-based de-noising scheme

Traditional filtering technique to remove high-frequency noises consists of low-pass and Butterworth filter [29]. The issue with it is that it dulls any rising/falling changes in the data which can be the CSI variation caused by water flow pattern. Low pass noise filter removes high-frequency signals generated in a complex indoor environment due to

18

electromagnetic interference, temperature changes and air pressure [1]. Here we apply the wavelet-based de-noising method to remove unwanted noise from the signal and smooth the CSI data after the outlier removal. It is better than the Fourier-based de-noising scheme in certain aspect since it captures both frequency and time domain information. Therefore, to conserve the transitional changes we adopt 2nd level wavelet based de-noising scheme [30]. We performed a level-2 wavelet transform on the CSI data after outlier removal and observed it much cleaner at the same time preserving the general CSI variation changes.

The basic idea behind the wavelet-based de-noising is that it localizes the feature in our data to different scales and can preserve important signals while removing the noise. In simpler terms, wavelet transform concentrates signal and image features in a large magnitude wavelet coefficient. Wavelet coefficient which is small in value represents noise and we can remove them without affecting the signal. As can be seen in Figure 7, wavelet-based de-noising scheme filters the original signal collected and uses a second level wavelet de-noising scheme to de-noise the signal preserving the transitional changes caused by the flow of water. As we saw, after the outlier removal using the Hampel identifier, there still remain some noisy components which need to be filtered out. And wavelet based de-noising technique is efficient in removing the remaining noisy components from the signal while preserving the variation due to the water flow.
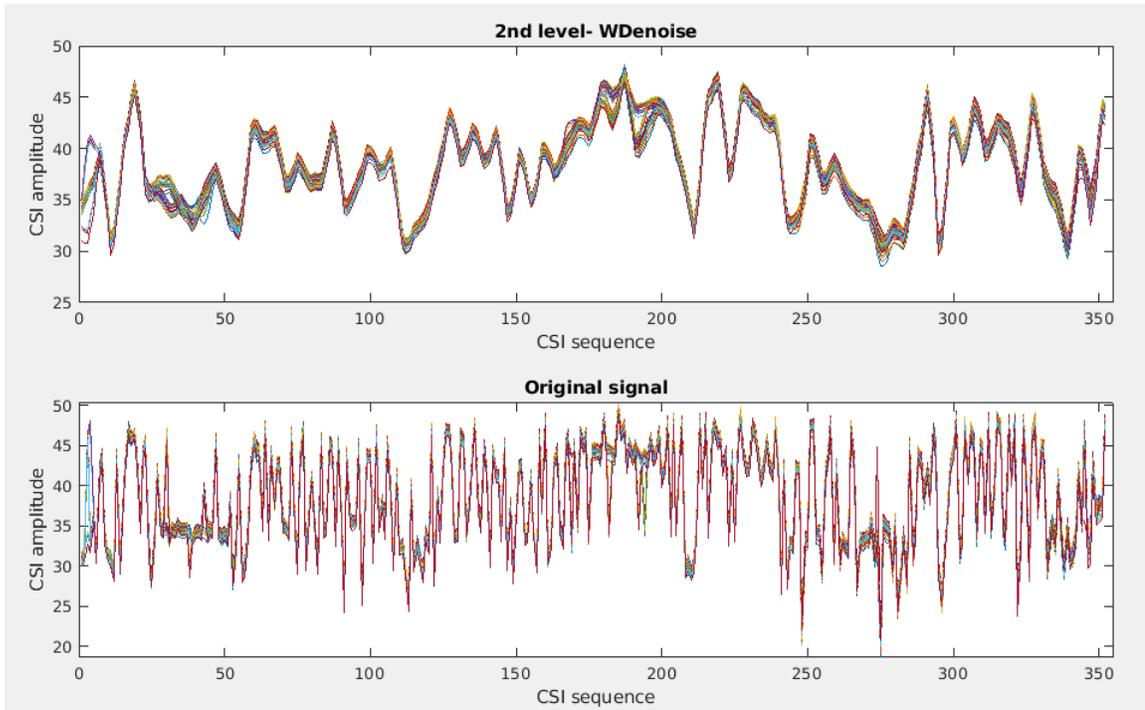
Figure 7 wavelet based de-noising scheme to filter the signal

3.3 PCA based dimension reduction

PCA (Principal component analysis) is a dimension reduction technique that reduces a large set of variables to a small set while still preserving most of the information from the large set. In other words, PCA is used to reduce data vectors described by n attributes or dimensions. PCA searches for *k* n-dimensional orthogonal vectors that can be used to best represent the data, here $k <= n$. In the below Figure 8 shows the first two principal components, as evident the two components, are orthogonal to each other. The basic procedure for performing PCA analysis on any data sets is:

1. Normalization of data, where the input data is normalized so each attribute falls within the same range. This ensures that attributes with large domains do not dominate attributes with the small domain.

2. Computation of k orthonormal vectors that provide a basis for normalized input data. These are unit vectors that each point in a direction perpendicular to the others and are referred to as the *principal components*.

3. The principal components are ordered in decreasing significance. The principal components basically create a new set of axes for the data.

   The first principal component shows the most variance among data, the second shows the second highest variance and so on.
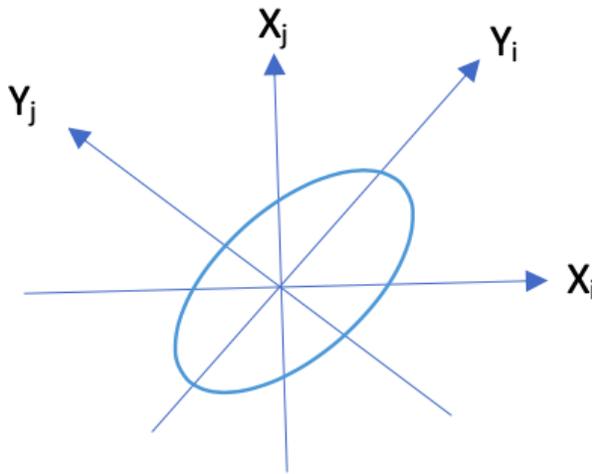
Figure 8 Principal component analysis. Here Y_i and Y_j are the first two principal components for the given data.

4. Most often, the first four principal components are enough to represent more than 90% of the data. Because the components are arranged in decreasing order of significance, the data size can be reduced by eliminating the weaker components.

PCA is being used in the various data mining process. It can handle data in various forms, ordered and unordered attribute. It can be used to reduce multidimensional data of more than two dimensions to two dimensions. PCA has shown to be better at handling sparse data as compared to wavelet transforms.

21

In our experiment, we compute PCA in four steps, after which we select the optimal number of principal components to represent our data:

i. **Standardize the data:** Usually, the first step in PCA includes standardizing the data by removing sample mean from each observation. This is done to remove any bias and it centers and scales the data.

ii. **Coefficients calculation:** Matlab is used for getting the coefficients, score and latent from the standardized data. The coefficient is also known as the loading for an n-by-p data matrix X. The coefficient matrix is p-by-p. PCA uses SVD (singular value decomposition) to get its data.

iii. **Eigen decomposition:** Eigen decomposition is used to find the score/ eigen vector of the dataset. In the matrix, the first column represents first PC and it contains the highest eigenvalue followed by 2nd component and so on.

iv. **Selecting the number of PCs:** This is calculated from the number of PCs used to represent a good variance of data. From our observation, the first four PCs are enough to represent 90% of the data.

From our experiment, we learned that the changes in CSI amplitude for each subcarrier in a Tx-Rx pair due to the flow of water from the shower head are correlated. Figure 9 and 10 plot the projected CSI amplitude of top 4 principal components. Unfortunately, the first principal component fluctuates severely regardless of any water flow. That's why we don't use it for further processing. After removing the first PC, we use the next four principal components for feature extraction [31].
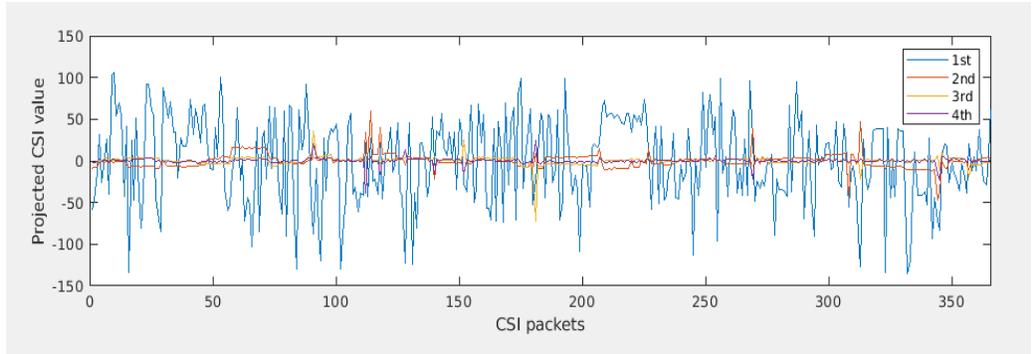
Figure 9 Projected CSI value of first 4 principal components.
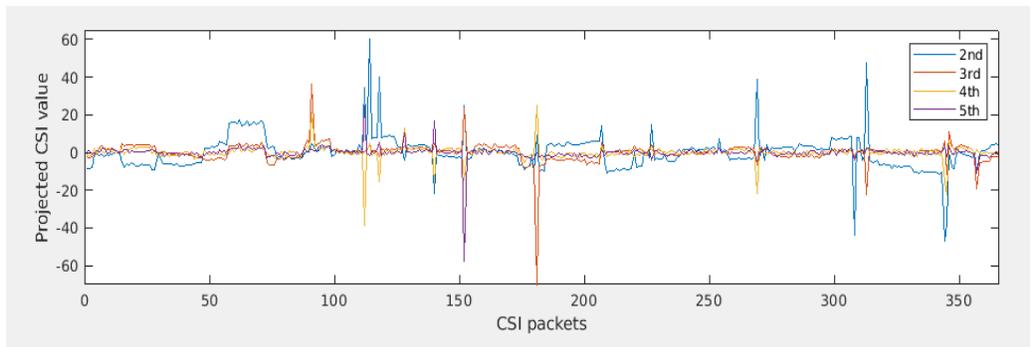
Figure 10 Projected CSI value after removal of first principal component.

3.4 Summary

In this chapter, we covered the CSI data pre-processing phase. The raw CSI data collected contains a lot of noise and for further processing, those noises needed to be removed. First, we used an outlier removal technique based on the Hampel identifier to get rid of the outliers. But still, there is noise present in the signal for that we use wavelet based de-noising scheme which removes the unwanted noise and preserves the variation in CSI representing the flow of water. At last, we use PCA based filtering on the de-noised CSI sequence to extract more representative signals. We discard the first principal component and use the next four principal components for feature extraction.

CHAPTER 4: CLASSIFICATION AND PERFORMANCE ANALYSIS

The objective of this chapter is to recognize the water flow and also the different patterns of the water flow using a multi-class classifier. We will first discuss the feature extracted from the PCA analysis, and the method used to get those features. Since we are planning to differentiate between water flowing at different rates, we used multi-class SVM classifier and discussed how we tuned the parameter to get the best results. After that, we discuss our data collection process as if how the CSI data was collected from the two routers when water was flowing from the shower head. We also assess our results that whether the features obtained from the Hampel identifier improve our results, by comparing the results with features obtained on signal collected from raw CSI data. At last, we present our results by highlighting the accuracy and cross-validation rate. We also discuss how our approach of collecting CSI data from the flowing water is really unique and a novel approach.

4.1 Feature Extraction

For the flowing water detection and pattern recognition, we extracted various features from the raw CSI data collected. The features collected are either time-domain or based on subcarrier correlations. For time-domain, the features are extracted from the de-noised CSI signals. Whereas the subcarrier correlation features are extracted from the PCA's four principal components.

4.1.1 Features extracted from Time

The following features are time-based:

1. Mean 2. Standard deviation 3. Median Absolute deviation 4. Maximum 5. Minimum 6. Skewness 7. Kurtosis

The above features are collected from each Tx-Rx pair. Since we are using 2*2 pair of

antennas, there is a total of 28 features (7*4). The formula for the above features are

discussed below:

1. Mean: Calculates mean of filtered CSI data

2. Standard deviation: used to measure the dispersion of a set of data values

$$\sigma = \sqrt{\frac{\sum(x-\overline{x})^2}{n}} \quad , \qquad \sigma = \text{lower case sigma} \qquad (4.1)$$

3. Median Absolute deviation(MAD): It is used to represent the median of the absolute

deviation from the data's median i.e.

$$\text{MAD} = \text{median } (|X_i - \tilde{X}|) \qquad (4.2)$$

4. Maximum takes the maximum magnitude value from the filtered CSI data.

5. Minimum takes the minimum value from the extracted CSI data.

6. Skewness is used to define the extent to which a distribution differs from a normal

distribution.

7. Kurtosis is a measure of the tailedness of the probability distribution of a real-valued

random variable. In other words, it tells the sharpness of the peak of a frequency-

distribution curve. It is a descriptor of the shape of a probability distribution.

4.1.2 Features extracted from subcarrier correlations

Apart from these features selected, there are also features selected from the principal

component analysis which uses an orthogonal transformation to convert a set of

observations of correlated variables into a set of linearly uncorrelated variables. The

model leverages our observation that the variation in CSI amplitude of all subcarriers due

to the flow of water are correlated. Therefore, it applies principal component analysis

(PCA) to the filtered subcarriers to extract the signal that only contains variation caused

by the flow of water. In this way, there is no need to include all the 56 subcarriers in the features as most of them are correlated. Only four principal components are enough to represent more than 90% of the CSI data. Since the first principal component contains unwanted noise, it was ignored during analysis.

4.2 Multi-Class Support Vector Machine (SVM) classification

To distinguish different settings of the water flow, a multi-class support vector machine [32]is applied. Support vector machine is capable of linear or non-linear classification. It is one of the most popular models in machine learning and is best suited for classification of complex but small or medium-sized datasets. In simpler terms, support vector machine is a discriminative classifier which uses a hyperplane to separate data of different labels. In linear classification, this line is a hyperplane which divides the dataset into two parts. As shown in Figure 11.
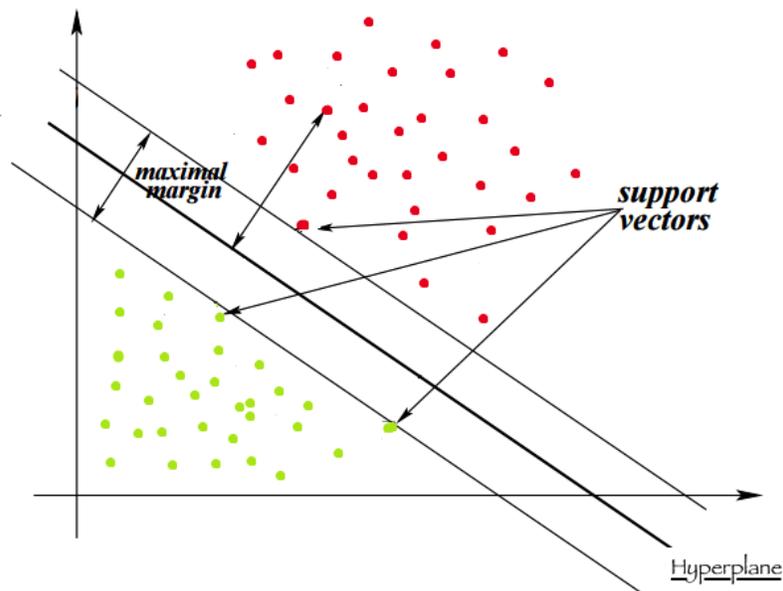

Figure 11 Support vector machine for linear classification

For solving non-linear classification problems, SVM uses kernel trick. Kernel trick greatly reduces the computational time. If provided with a training dataset $D = \{(x_1, y_1), (x_2, y_2),\ldots,(x_n, y_n)\}$, where x_i represents the feature vector and y_i represents the class, soft margin SVM algorithm, introduced by Vapnik et al. in 1995 solves the following optimization problem

$$\min \quad \frac{1}{2} \vec{w}^T \vec{w} + C \sum_{i=1}^{n} \xi_i$$

$$\text{for} \quad y_i (\vec{w}^T \cdot x_i + b) \geq 1 - \xi_i, \qquad \forall(x_i, y_i) \in D \qquad (4.3)$$

$$w, b, \text{ and } \xi_i \geq 0$$

Where ξ_i measures the degree of misclassification of the data point. Higher the value of ξ_i leads to more bias and less variance. Here, C known as the cost of misclassification determines the tradeoff between the margin size and amount of error in training. Lagrange multipliers can be used to solve the above equation to get the hyperplane separating the classes.

In other words, in the case of non-linear or multi-class SVM the following steps are followed:

i) The input feature vector (mean, STD, and others) is transformed into higher dimensional space which allows it to separate the training data linearly.

ii) Soft margin SVM is used to find the hyperplane with the maximal margin, which uses hinge loss function. *kernel trick* is introduced which basically computes the dot product between feature vectors as if they have been transformed into a higher dimensional space, without actually transforming any of the vectors, illustrated in the below Figure 12.
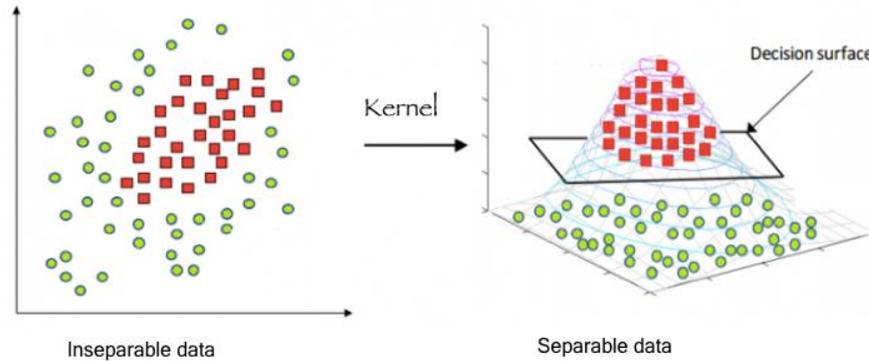
Figure 12 Kernel trick to separate inseparable data

There are different kernels available, such as the polynomial kernel, Gaussian kernel, Gaussian radial basis function (RBF) and many others. The kernel used in our SVM classifier is RBF [33]. It is a general-purpose kernel and is used when there is no prior knowledge about the data. It has been proven to be a good choice by providing satisfactory results. The equation for the RBF kernel is

$$K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2) \tag{4.4}$$

Here γ (gamma) is the kernel parameter, a large value of gamma leads to higher bias and lower variance, and vice versa.

4.2.1 Cross-validation and grid-search

Now, we have two parameters C and gamma whose values need to be tuned. Most often grid search is used to find the optimum values of C and gamma. GridSearchCV library in python [34] can be used to search for the hyper parameter, which evaluates all the possible values of the hyper parameter using cross-validation technique.

Cross-validation is done to assess the predictive performance of the model and to check how it performs on a new unseen dataset. The data is split into training, validation and testing parts. The classifier is trained on the training data set and then tested on the

28

test set. The concept is to keep the training and testing data separate so that the classifier can be tested on unseen data. 10-fold cross-validation is done on the training data to prevent *overfitting,* which occurs when you model the training data too well such that it does not perform well on the testing data. It usually occurs when the sample size is small, and the number of features is large. In cross-validation technique the training set is divided into k-fold, 10 in our case and each instance of the training set is predicted giving a cross-validation accuracy.

Grid search was performed to compute the optimum values of C and gamma. The one which resulted in the highest cross-validation accuracy was selected. In our scenario, after the grid search the optimal values for C and gamma came as $C = 100$ $\gamma = 0.001$. The SVM classifier was built, trained and tested in python with the help of 'Scikit-learn' package [34]. It features various classification, regression and clustering algorithms including support vector machine. We used the model selection technique for grid search and cross-validation modules.

4.3 Dataset collection

In this section, we would go through the process involved in the collection of data and the environmental settings used.

4.3.1 Experimental setting

In our experiment, we used two TP-link routers of model TL-1043ND v3.0 at both the transmitter and receiver sides. We installed the Atheros CSI tool [14] at both the routers. Before installing the Atheros CSI tool, we upgraded the router's firmware to OpenWRT as it supports the Atheros CSI tool. The CSI data was collected between the two transmitting and receiving antennas. So, in total there are 4 TR links (2*2). The two

routers run on 2.4 GHz and use channel 6 as it proved to be the one with the less

interference. The routers are connected with a laptop using a LAN interface on both

sides, from which we give the commands for sending and receiving the CSI data as can

be seen from the below Figure 13.

Figure 13 Experimental setting for CSI data collection using two routers.

For each packet, there are 2*2*56 CSI complex series data. As discussed in section 2.2

the phase data is very irregular due to phase frequency offsets. So, here we only use the

amplitude data. For feature extraction, the amplitude of the 56 subcarriers from each

packet was used to calculate the standard deviation, mean and other features. We utilize

the PCA (principal component analysis) to reduce the number of correlated subcarriers

from 56 subcarriers to 4 principal components per link, as those are enough to represent

more than 90% of the data. And since we used 4 TR links, the total number of PCs are

4 PCs * 4 links = 16 PCs.

To give commands to the routers for sending and receiving CSI data, two laptops

are used. Both routers are configured with sending and receiving CSI commands. So, any

one of them can be used as a sender or receiver. The two routers are connected in a

Master and Client structure, with the help of the *luci* web interface which is interfaced in the OpenWRT router.

4.3.2 CSI Data collection

The CSI data has been used for various purposes like human detection [11], keystroke recognition [3], gait recognition [2] and various other purposes. Most of those experiments are done in a lab where they set up two routers or laptops at two ends and a person walks in between. However, our scenario is entirely different. Since we test the variation in CSI data due to flowing water from a shower head, the experiment cannot be performed in the lab. In this work, the routers and laptops are set up at home in the bathroom. From Figure 14 we can see that one router is kept inside the bathtub at one end, whereas the other router is kept outside on the chair at a distance of around 2 meters.

Figure 14 CSI data collection setting, showing the shower head, router and the laptop used.

During the data collection procedures, the water flowed from the shower head at different

settings between the two routers i.e. in the Line of sight scenario. The CSI data was

collected for a total of 7 settings, 6 settings included in the shower head and one extra

when there was no water flowing. For each setting a different number of packets are

collected. The packets are collected for an approximate period of 50 seconds. Here, we

refer to one packet as the CSI data from one timestamp which contains the CSI complex

value in the format of 2*2*56 CSI data. Table 4.1 mentions the total number of packets

collected for each setting. We call the setting when no water is flowing as still. All the

other settings are numbered from 1 to 6. The total number of packets collected for all the

settings amount to 2795.

Table 4.1 Number of packets collected for each setting.

SETTINGS	NUMBER OF PACKETS
STILL	635
SETTING1	321
SETTING2	315
SETTING3	299
SETTING4	362
SETTING5	379
SETTING6	484
TOTAL	2795

The flow rate for the water from the shower was at 1.8 gpm (gallons per minute). To

capture the variation in CSI data due to flowing water we send packets from the sender at

a rate of 15 packets per second, which proved to be enough to detect the CSI variation

due to flowing water. For each setting the CSI data was collected for around 50 seconds.

4.4 Experimental Results

In this section, we will show the results obtained by the multi-class SVM classifiers for

the collected CSI data. We will evaluate the results obtained from the SVM classifier and

check the cross-validation accuracy. Also, to better visualize the result we will draw a confusion matrix.

Since there is not any previous research work which evaluated the CSI data variance due to water flow, it is difficult to compare the accuracy obtained by our multi-class SVM classifier with other works. Nevertheless, the results obtained by our classifier seem quite promising. The 10-fold cross-validation accuracy comes at *90.43%* obtained on the training dataset. The model is trained on 9 folds as the training data and validated on the remaining data. The final accuracy obtained is the average of the accuracy obtained in each fold. Cross-validation is used to detect *overfitting* which occurs when the model is too closely fit on the training dataset such that it does not perform well on unseen data. If the cross-validation accuracy is too high on the training dataset and performs poorly on the test dataset, in that case, the model is said to be over-fit. In our case, the cross-validation accuracy is slightly less than the accuracy on the test data, which is the ideal case.

Grid search was done to find the optimum value of cost and gamma for the RBF kernel used in the multi-class SVM classifier. Using Scikit-Learn's GridSearchCV library in python, we just need to provide the hyper parameters and the range of values to try out. In our scenario, the optimum value for C came out to be 100 and gamma 0.001 with a score of 0.93.

The accuracy of the trained multi-class SVM classifier on test data came at *90.35%.* Which is slightly less than the cross-validation accuracy of *90.43%* on the training data. It indicates our model is robust and performs well even on unseen data and there is no problem of *overfitting*. As can be seen in Figure 15, the confusion matrix

shows the accuracy of the classifier, in simple terms confusion matrix tells the number of times the classifier confused one setting with the other. The rows represent an actual class, while each column represents a predicted class. The figure below shows the actual number which the classifier predicted for all the settings and is a good visual representation of the confusion matrix using mat plot library in python.
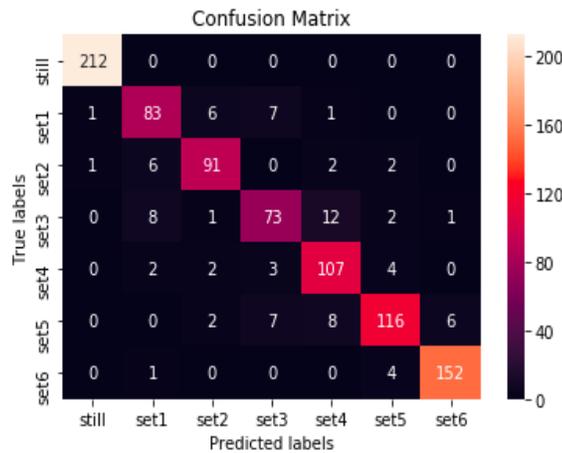
Figure 15 Showing Confusion matrix to evaluate the performance of the classifier.

In the confusion matrix the first label is for still, then set1 is for setting 1 and so on till set6. In total there are 7 settings, and from the array matrix, we can see how many features were wrongly classified in the predicted class from each column. In Figure 16 we highlight the wrong features. As can be seen from the figure the column of class 4 and 5 are relatively bright, meaning many features from class 4 and 5 got misclassified. One of the reasons could be the close resemblance of setting 4 with setting 5 which led to high misclassification of the label.
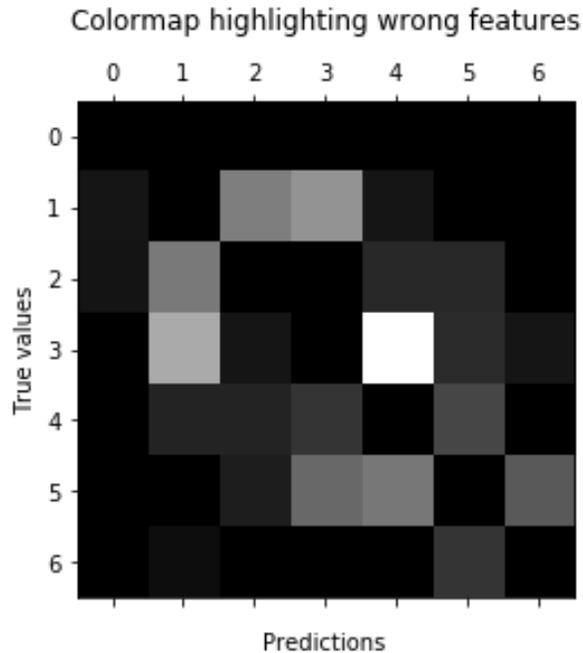
Figure 16 Confusion matrix Highlighting the wrong features

4.5 Summary

In this chapter, we presented our results obtained from the multi-class SVM classifier for different settings of the water flow. We explained all the features extracted from the filtered CSI data and the process used to obtain them. We discussed the principal component analysis as to how it is being used here for dimensionality reduction. Since most of the subcarriers from one Tx-Rx pair are correlated we used PCA to obtain the principal components from the 56 subcarrier CSI data and selected four principal components which best represent the CSI data. We described the scenario in which the data was captured, which is quite different from the previous research work done. As our experiment required flowing water it couldn't be performed in the lab and has to be done at home. The number of packets collected for each setting was mentioned.

We discussed the multi-class SVM classifier and the RBF kernel trick used in it. The C and γ parameters are optimized using the grid-search technique. Cross-validation was done to make sure there was no *overfitting* in the data. At last, evaluated the final accuracy of the model which came above *90%.* In short, it can be said that this is the first work which leverages the CSI physical layer information for detecting the flow of water from the shower head and is also very cost effective as compared to other devices such as USRP radio.

CHAPTER 5: CONCLUSION AND FUTURE WORKS

In this chapter, we will summarize our research work and look at some of the future prospects where it could be useful. Wi-Fi based detection of flowing water pattern is a novel approach which can have multiple uses. It can be utilized by firefighters to measure the effectiveness of their operation; it can be used by farmers for irrigation purposes. Moreover, it can also be used in outdoor conditions to detect rain.

Wi-Fi based flowing water detection is implemented with the help of two routers running Atheros CSI tool on 802.11n NICs and the commands are given to the routers through a laptop which creates an SSH tunnel to the router. The experiments were conducted at home under normal condition.

5.1 Conclusion

In Chapter 1, a basic introduction was made to the Wi-Fi system and we discussed how it is better in activity recognition than video surveillance or wearable devices. We went through some of the related works where Wi-Fi signal is used to for fall detection, gait recognition, keystroke recognition, and various other methods. Our approach of using the Wi-Fi signal for detecting the flow of water is quite unique and to the best of our knowledge, we couldn't find any relevant work regarding that. In chapter 1 we also went through the technical challenges faced during our research work. There were many challenges faced which helped me to learn a lot from them and made me more open toward new challenges.

Chapter 2 covered the literature review and a basic introduction to channel state information (CSI). Channel state information has gained a lot of attention in recent years because of its fine-grained information and resilience towards outside environment

changes. After that, in the wireless signal processing section, we discussed scattering, reflection and diffraction of Wi-Fi signals and how it affects our experiment. MIMO system was covered as we are using MIMO technology here to send the CSI data from two transmitting antennas to two receiving antennas. MIMO uses spatial multiplexing to transmit different bits of information in parallel channels which increase the data rate. Then we made a brief introduction to the channel state information and how it utilizes scattering, fading and power decay of signal to predict changes in the environment. We went through the process required to capture the channel state information (CSI) from the NIC (network interface card). We also discussed the Atheros CSI tool installation process, and the challenges faced by it. Atheros CSI tool requires specific hardware on which it can be installed.

Chapter 3 features the CSI data pre-processing stage and is the most important stage. Without it the raw CSI data will contain random noises, making it unusable. Since the CSI captures the changes in the environment, there are various unwanted variations recorded which must be filtered out. Here, we use Hampel-identifier for outlier removal. The raw CSI data containing burst noises which generate outliers are removed with the help of Hampel-identifier. Even after that, there is still some noise left in the CSI signals, for that we use wavelet based de-noising scheme. Wavelet de-noising technique removes the unwanted noise while preserving the natural CSI variations. PCA was used for dimensionality reduction. We used four principal components obtained from PCA to represents the 56 subcarriers.

Chapter 4 deals with the feature extraction process and performance analysis. Features are extracted from the filtered CSI signal after the outlier and random noise are

removed. There is a total of 11 features extracted from each Tx-Rx pair. After that, we discussed the classifier used, i.e. multi-class SVM classifier with k-fold cross-validation technique [35]. Then we discussed the grid-search technique to find the optimum value of classifier hyper parameters. The SVM classifier yielded an accuracy of 90.35%, with a cross-validation accuracy of 90.43%, which proves there was no over-fitting in the dataset. This chapter also covers the process involved in the dataset collection which is a bit different as compared to the data collection process for human detection technique.

5.2 Future Works

In this experiment, our work is performed at home which is very much restricted in space. In future work, we would like to test it in an open environment where it could be used by firefighters to measure the effectiveness of their operation. Under open conditions, we would have to consider the effects due to various human movements and accordingly apply filtering technique to capture the desired CSI variations.

In chapter 4 we utilized the multi-class SVM classifier to differentiate between different settings of the shower head. The motivation for using SVM classifier was that it has been used in various previous baseline works. It would be interesting to measure the accuracy using other methods such as Extreme Learning Machine [36] which has lower training time, Random forest [37] which operates by constructing a multitude of decision trees at the training time. Artificial Neural Network [38] is another method which seems quite promising. In Artificial Neural Network the online training is quite simple as compared to online SVM fitting with quick prediction time.

Currently, our work mainly concerns with the water flowing from a shower head at a given setting for a fixed period of time. However, in real-time people keep changing

the setting from one to another. We could develop an algorithm to segment the CSI data

collected based on different settings and then better predict the setting in real-time. As of

right now, the data is collected and processed off-line. It would be great if a real-time

detection can be done on the flowing water which could serve a much more practical

purpose.

BIBLIOGRAPHY

[1]   H. Zhu et al., "Robust and passive motion detection with COTS Wi-Fi devices," *IEEE Tsinghua Sci. Technol. J., vol. 22, no. 4*, pp. 345–359, 2017.

[2]   W. Wang, A. Liu, and M. Shahzad, "Gait recognition using wifi signals", *UbiComp'16 ACM conf. on Pervasive and Ubiquitous Computing,* pp. 363-373, 2016.

[3]   K. Ali, A. Liu, W. Wang, M. Shahzad, "Keystroke Recognition using WiFi Signals", *21st Annual Conf. Mobile Computing and Networking, MobiCom'15*, pp. 90-102, 2015.

[4]   H. Abdelnasser, M. Youssef, and Khaled A. Harras, "WiGest: A ubiquitous WiFi- based gesture recognition system", *IEEE Conference on Computer Communications (INFOCOM)*, 2015.

[5]   E. Martin, O. Vinyals, G. Friedland, and R. Bajcsy, "Precise indoor localization using smart phones", *Proceedings of 18th ACM international conf. on Multimedia,* pp. 787-790, 2010.

[6]   S. Sigg, U. Blanke, and T. Gerhard, "The telepathic phone: Frictionless activity recognition from Wi-Fi RSSI", *IEEE International conf. on Pervasive Computing and Communications(PerCom),* 2014.

[7]   F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3d tracking via body radio reflections", *11th USENIX Symposium on Networked Systems Design and Implementation,* pp. 317–329, 2014.

[8]   B. Kellogg, V. Talla, and S. Gollakota, "Bringing Gesture Recognition to All Devices", *11th USENIX Symposium on Networked Systems Design and Implementation,* 2014.

[9]   X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach", *IEEE Transactions on Vehicular Technology, vol. 66, Issue 1,* pp. 763-776, 2017.

[10]   C. Han, K. Wu, Y. Wang, L.M. Ni, "WiFall: Device-free fall detection by wireless networks", *IEEE Infocom, Conference on Computer Communications*, pp. 217-279, 2014.

[11]   W. Xi et al., "Electronic frog eye: Counting crowd using WiFi", *IEEE Infocom, Conference on Computer Communications*,2014.

[12]   C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-Invasive Detection of Moving and Stationary Human with WiFi", *IEEE J. Sel. Areas Commun., vol.33, no. 11,* pp. 2329-2342, 2015.

[13]   D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool Release: Gathering 802.11n Traces with Channel State Information", *ACM Sigcomm*, 2011.

[14]   Y. Xie, Z. Li, and Mo Li, "Precise Power Delay Profiling with Commodity WiFi", *21st Annual Intl. Conf. on Mobile Computing and Networking (MobiCom '15) ACM*, 2015.

[15]   Chowdhury, Tahmid Z. "Using Wi-Fi Channel State Information (CSI) for Human Activity Recognition and Fall Detection." T. University of British Columbia, 2018. Web. 6 May 2019. https://open.library.ubc.ca/collections/ubctheses/24/items/1.0365967. Electronic Theses and Dissertations (ETDs) 2008+.

[16]   J. Winters, "On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment", *IEEE Journal Selected Areas in Commun., vol.5 no. 5,* pp. 871-878, 1987.

[17]   L. Lu et al., "An Overview of Massive MIMO: Benefits and Challenges", *IEEE Journal Selected Topics in Signal Processing, vol. 8 no. 5,* pp. 742-758, 2014.

[18]   C. Oestges and B. Clerckx, "MIMO Wireless Communications: From Real-World Propagation to Space-Time Code Design", New York: Academic Press, 2007.

[19]   T. Wang, W. Heinzelman and A. Seyedi, "Minimization of transceiver energy consumption in wireless sensor networks with AWGN channels", *IEEE 46th Annual Conf. Commun., Control and Computing,* 2008.

[20]   W. Wang, et al., "Understanding and Modeling of WiFi Signal Based Human Activity Recognition", *21st Annual conf. on Mobile Computing and Networking (MobiCom '15) ACM,* pp.65-76, 2015.

[21]   Z. Yang, Z. Zhou and Y. Liu, "From RSSI to CSI: Indoor localization via channel response", *ACM Comp. Surv. vol. 46 no. 2, Article. 25*, 2013.

[22]   R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission", *Bell Syst. Tech. J., vol. 45,* pp. 1775-1796, 1966.

[23]   J.A.C. Bingham, "Multicarrier modulation for data transmission", *IEEE communication Mag., vol.28, no.5*, 1990.

[24]   H. Zhu et al., "R-PMD: robust passive motion detection using PHY information with MIMO", *IEEE 34th Intl. Performance Comp. Comm. Conf.,* 2015.

[25] W. Wang et al., "Device-Free Human Activity Recognition Using Commercial WiFi Devices", *IEEE journal selected areas in Commun.,* 2017.

[26] Davies, L. and Gather, U., "The identification of multiple outliers", *Journal of the American Statistical Association 88,* 1993.

[27] F R. Hampel, "The Influence Curve and its Role in Robust Estimation", *Journal of American Statistical Association, vol. 69, no. 346,* 1974.

[28] C. Becker and U. Gather, "The Masking Breakdown Point of Multivariate Outlier Identification Rules", *American Statistical Assoc., vol. 94, no. 447, pp. 947-955*, 1997

[29] J. Zhang, et al., "WiFi-ID: Human Identification Using WiFi Signal", *Intl. conf. Distr. Computing Sensor Systems*, 2016.

[30] W. He, K. Wu, Y. Zou and Z. Ming, "WiG: WiFi-Based Gesture Recognition System", *Intl. Conf. Computer Comm. Network*, 2015

[31] J. Han and M. Kamber, "Data Mining: Concepts and Techniques, Morgan Kaufmann", San Francisco, Calif, USA, 2001.

[32] A. Géron, "Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems", Sebastopol, CA, USA: O'Reilly, 2017.

[33] M. Girolami, "Mercer kernel-based clustering in feature space", *IEEE Transactions on Neural Networks, vol. 13, no. 3*, pp.780-784, 2002.

[34] F. Pedregosa et al. "Scikit-learn: Machine learning in Python", *Journal of Machine Learning Research, 12:2825–2830*, 2011.

[35] Soaresda Silva, B., Teodorolaureano, G., Abdallah, A. S., & Vieiracardoso, K. (2018). "WiDMove: Sensing Movement Direction Using IEEE 802.11n Interfaces", *In 2018 IEEE Canadian Conference on Electrical and Computer Engineering, CCECE 2018 (Vol. 2018-May).* [8447627] Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CCECE.2018.8447627.

[36] H. Zhou, B. Huang, X. Lu, H. Jiang, and L. Xie, "A Robust Indoor Positioning System Based on Procrustes Analysis and Weighted Extreme Learning Machine", *IEEE Trans. Wireless Comm., vol. 15, no. 2*, 2016.

[37] H. Zhou, et al., "FreeDetector: Device-Free Occupancy Detection with Commodity WiFi", *IEEE Intl. Conf. Sensing Comm. Networking,* 2017.

[38]   Yao, X. "Evolving artificial neural networks", *Proceedings of the IEEE, 87(9):1423–1447,* 1999.