



SAPIENZA
UNIVERSITÀ DI ROMA

A Robust Person Identification Method Through Wi-Fi Signals

Faculty of Information Engineering, Informatics, and Statistics
Applied Computer Science and Artificial Intelligence

Francesco Danese
ID number 1926188

Advisor
Prof. Luigi Cinque

Co-Advisors
Prof. Danilo Avola
Prof. Daniele Pannone

Academic Year 2022/2023

Thesis not yet defended

A Robust Person Identification Method Through Wi-Fi Signals
Sapienza University of Rome

© 2023 Francesco Danese. All rights reserved

This thesis has been typeset by \LaTeX and the Sapthesis class.

Author's email: francescodanese01@gmail.com

Abstract

This thesis deals with myself.

Contents

1	Introduction	1
2	The wireless channel	3
2.1	Electromagnetic waves	3
2.2	I/Q sampling	5
2.3	Orthogonal Frequency-Division Multiplexing	6
2.4	Channel State Information	7
3	Data preprocessing	8

Chapter 1

Introduction

In the last years, Wi-Fi technology is turning into a ubiquitous and fundamental tool, not only for providing constant and easy-access internet connection but also for enabling the development of the Internet of Things and radio-based sensing applications. Cutting-edge Wi-Fi sensing systems use wireless signals to detect events or changes, such as motion and gesture recognition [1], in the permeated environment leveraging the disturbances and variations in the received radio signal. In this work, Wi-Fi sensing technology is employed for person Identification, a task usually achieved by vision-based systems through the use of RGB cameras. In common scenarios, person Identification is a classification task where the input features, such as an image or a video of a particular individual, is examined in order to recognize him among a specific set of people. To reach this goal, in modern approaches the raw input image is analyzed in order to model a person appearance and physical constitution, by extracting some key features that will be use for inference, as shown in Fig. 1.1. State-of-the-art methods include leveraging deep learning models trained to identify a human, like: Convolutional Neural Network for detecting the face of the person [2] and recognize him; Recurrent Neural Networks [3] and their variants, such as LSTMs, for working with sequences of data like video sequences and exploiting the unique person behavior like the walking gait [4]; Attention mechanisms, as seen in models like Transformers [5], for their ability to focus on relevant parts of an image or sequence therefore locating the discriminative features. Although vision-based methods have achieved remarkable results, several issues arise during real-world scenarios, including: subject occlusion and pose; different viewing angles and image scale; ambient changes in illumination; modifications in personal look (clothes, haircut, glasses) and privacy concerns due to the sensitive nature of personal imaging data. To address these challenges, a non-conventional approach can be adopted, that walks away from images and vision avoiding the aforementioned issues, through the use of completely different types of data: Wi-Fi signals.

Wi-Fi technology has revolutionized the way we connect and communicate in the digital age. It enables wireless communication between devices by utilizing radio signals transmitted by access points (APs). The signals can be influenced by various factors, such as objects and people, leading to signal variations along their path. These variations, essential for understanding the wireless environment, are captured through measurements like the Received Signal Strength Indicator (RSSI) and



Figure 1.1. Image-based person ID within a group of 5 people.

Channel State Information (CSI), which play a pivotal role in implementing wireless sensing applications. The RSSI serves as a measure of the reduction in the strength of radio signals as they travel through space, capturing the way signals diminish over distance. Even though this metric has found widespread utilization within numerous indoor localization systems [6, 7], which goal is to find where a certain target is in an enclosed location with meter-precision, it experiences unsustainable performance drops in complex situations, due to several limitations. The problems associated with RSSI stem from its simplicity and lack of granularity, which restrict its ability to capture detailed signal characteristics and changes accurately. In fact, for a given wireless data packet, RSSI provides a single value computed at MAC layer, that represents the overall signal strength, without differentiating between the effects of multipath interference, signal reflections, or other external factors. Moreover, the relationship between RSSI values and the real signal strength is often non-linear, so that small changes in RSSI may not correspond to proportional changes in signal strength, making it challenging to interpret signal variations accurately. For these relevant reasons, RSSI fails to provide detailed insights into how the signal interacts with the surrounding environment [8]. On the other hand, the CSI is derived from the PHY layer's orthogonal frequency-division multiplexing (OFDM) transmission technology and includes detailed signal information defined at the subcarrier level, making it more stable and capable of carrying a higher level of signal information. CSI measurements offer access to a broader range of signal properties, including amplitude and phase for each subcarrier, providing higher granularity than RSSI [9]. Furthermore, CSI can be computed using commodity Wi-Fi devices, making it a popular choice in Wireless sensing applications. To ensure the quality of received CSI measurements, signal pre-processing techniques are often employed, such as removing amplitude outliers and phase offsets, and machine learning or deep learning approaches are subsequently applied to address the specific final task.

Ultimately, previous investigations on electromagnetic waves interactions with biological tissues [10] have revealed that signal propagation around a human body is highly dependent on biometric characteristics, such as body structure, weight, height, skin and tissues composition, with a great variability among different individuals. The aim of this work is then to harness the biological-related radio signals fluctuations, captured by CSI, that occur in a Wi-Fi-pervaded environment when a person is present, and extract important insights to enable the person Identification.

Chapter 2

The wireless channel

This section introduces the theoretical concepts related to the wireless communication channel, that are necessary to grasp the foundational technology behind Wi-Fi sensing applications, such as Human Re-ID

2.1 Electromagnetic waves

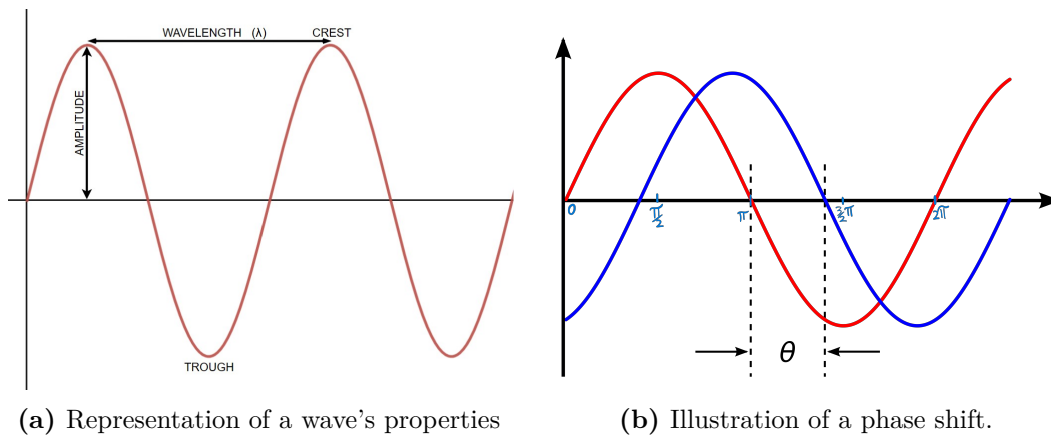
The fundamental characteristic of all wireless communications, including Wi-Fi transmission, is the use of Electromagnetic (EM) radiation to carry information, encapsulating data in the form of EM waves. In a vacuum, electromagnetic waves travel at the speed of light (around 299 792 km/s) denoted by the letter c . When passing through a material medium, they are slowed depending on the medium's permeability and permittivity but the air in the Earth's atmosphere is thin enough that EM waves travel very close to the speed of light. Each wave is produced by a transmitting antenna (TX) that generates synchronous oscillations between electric and magnetic fields, and it is characterized by several properties, shown in Fig. 2.1a: The wavelength, measured in meters, which is the distance between consecutive corresponding points of the same phase on the wave, such as two adjacent crests (the highest points), troughs (the lowest points), or zero crossings; The amplitude, in meters, that is the distance from the resting position to the maximum displacement of the wave; The frequency, measured in Hertz (Hz), that is the number of oscillation cycles per second or how many crests (or troughs) pass a location in a certain amount of time. The frequency is equal to $1/p$, where p is the period of the wave, which is the time in seconds that the signal takes to return to the equilibrium position, starting from the previous one (also called a full "cycle"). The mathematical link between the frequency f , the wavelength λ , and the wave propagation velocity v is expressed as follows:

$$f = \frac{v}{\lambda}$$

Where in vacuum $v = c$, the speed of light, as explained before. From the previous equation we get:

$$v = f\lambda \quad \text{and} \quad \lambda = \frac{v}{f}$$

indicating how the frequency and wavelength of an EM wave are directly proportional to the propagation velocity and inversely proportional to each other.



(a) Representation of a wave's properties

(b) Illustration of a phase shift.

The last key component used to describe a wave is the phase. The phase of a wave refers to its position within one complete cycle, describing how much of the cycle has occurred at a specific point in space and time. It is often measured in radians or degrees, and it helps determine where a wave is in its oscillation at a given moment. A phase angle of 0 represents the starting point of a cycle, while 2π (or 360°) corresponds to one complete cycle. When comparing two waves, the phase difference, or phase shift, is the angular difference between their phase angles, and it indicates how "out of sync" or "in sync" they are, as seen in Fig. 2.1b. Considering two sinusoidal waves, when the phase difference is a multiple of 2π , the addition of the two waves results in a constructive interference, leading to reinforcement and a larger amplitude. Vice versa, destructive interference occurs when the phase difference is an odd multiple of π , leading to cancellation and a smaller amplitude.

With those characteristic we can formulate the mathematical representation of a periodic wave, that tells us the momentary amplitude at instant t :

$$a(t) = A \cos(2\pi f t + \phi)$$

made up of the following components:

- $a(t)$: The instantaneous amplitude of the wave at time t .
- A : The amplitude of the wave, its maximum value.
- $2\pi f$: The angular frequency in rad/s , where f is the frequency in Hertz ($1/s$)
- t : The time in seconds, indicating the instant at which you evaluate $a(t)$
- ϕ : The phase offset, in radians, that determines the wave' starting position.

When ϕ is non-zero, the entire waveform appears to be shifted in time by $\frac{\phi}{2\pi f}$ seconds. A negative value represents a delay, and a positive value represents an advance. The above equation describes a sinusoidal wave as a function of time, and it's a fundamental building block in understanding and modeling periodic oscillations and waveforms in various fields of science.

2.2 I/Q sampling

In many real-world applications, information is transmitted through analog signals, which vary continuously over time. To process and analyze these signals using digital electronics, they need to be converted into a discrete digital form, a process known as sampling. I/Q sampling is a method used to sample and represent a complex-valued analog signal, from which we can retrieve both amplitude and phase information. The "I" component represents the in-phase or real part of the signal, while the "Q" component represents the quadrature or imaginary part of the signal. Denoting with A the amplitude and with ϕ_t the phase of the wave at instant t , we got:

$$I(t) = A \cos(\phi_t) \quad \text{and} \quad Q(t) = A \sin(\phi_t)$$

Now, we can construct the complex number at any given time t by combining the "I" and "Q" components

$$Z(t) = I(t) + jQ(t)$$

Where j is the imaginary unit ($j^2 = -1$). Therefore, as depicted in Fig. 2.2, any sinusoidal wave can be described by a composition of those two components (that is, by a single complex number), from which it is possible to obtain the amplitude and the phase using the following equations:

$$A = \sqrt{I(t)^2 + Q(t)^2} \quad \text{and} \quad \phi_t = \text{atan2}(I(t), Q(t))$$

The first formula is simply the Pythagorean theorem applied to the two sides, in our case I and Q, to find the hypotenuse, that matches the amplitude. The second one uses $\text{atan2}(x, y)$, a function that returns the angle formed by any vector $[x, y]$ with the positive x-axis, bounded in the interval $(-\pi, \pi]$. The angle is positive if counterclockwise, negative if clockwise. Additionally, since $\cos(x) + j \sin(x) = e^{jx}$, we can represent the I/Q sample in the Euler's form:

$$Z(t) = A(\cos(\phi_t) + j \sin(\phi_t)) = Ae^{j\phi_t}$$



Figure 2.2. Decomposition of a wave in I/Q components.

2.3 Orthogonal Frequency-Division Multiplexing

Orthogonal frequency-division multiplexing (OFDM) is a modulation scheme used in wireless communication systems such as 802.11n which encodes data streams into multiple overlapping subcarrier frequencies. In a traditional single-channel modulation scheme, each data bit is sent serially or sequentially one after another. With OFDM a single information stream is split among several closely spaced narrowband subchannel frequencies instead of a single Wideband channel [11]. This enables each substream's data rate to be lower than would be required by a single stream of similar bandwidth and makes the system less susceptible to corruption. Even though the subchannel frequencies are closely spaced and do overlap to a degree, they remain orthogonal, maintaining their distinct separation due to meticulous selection. This strategy ensures that any interference is effectively negated.

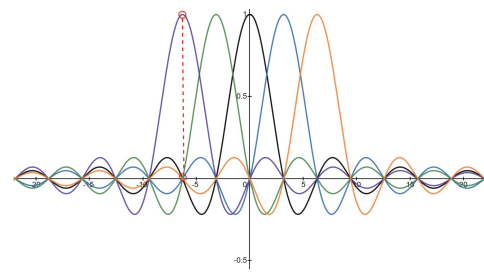
OFDM signals are modelled in the frequency domain as the *sinc* function:

$$\text{sinc}(f) = \frac{\sin(f)}{f}$$

The *sinc* function has a central peak at zero frequency and then oscillates symmetrically around zero. It is characterized by its main lobe, which contains most of its energy, and smaller side lobes that decrease in amplitude as you move away from the central peak, as seen in Fig. 2.3a. The choice of the *sinc* function is made due to its property of enabling orthogonal placement of each subcarrier with respect to the others [12], as illustrated in Fig. 2.3b where five *sinc* functions are strategically positioned so that the primary subcarrier frequency (one is marked in red) coincides with a frequency at which all accompanying *sinc* functions assume a value of zero. This orthogonal configuration ensures that the summation of all five selected subcarriers maintains its peak values, as presented in Fig. 2.3c.



(a) The *sinc* function for a single subcarrier signal.



(b) Orthogonal *sinc* functions for multiple subcarriers.



(c) Signals summation.

In order to transmit bits, a signal undergoes modulation. In wireless communication, modulation is the process of altering one or more characteristics of a periodic waveform, called the carrier signal, with a separate signal called the modulation signal that typically contains information to be transmitted. With OFDM, each subcarrier represents a single digital symbol that is being transmitted and can be modulated through various methods [13] such as Amplitude Shift Keying (ASK), Binary Phase-Shift Keying (BPSK) or Quadrature Amplitude Modulation (QAM). ASK involves switching the amplitude of the carrier signal between two different levels (high and low) to represent binary data while BPSK achieves the same thing by altering the phase (usually 0° and 180°) to represent the travelling bits. QAM, widely used in Wi-Fi communication systems, combines instead both amplitude and phase modulation to convey multiple bits of digital information in each modulation symbol, resulting in a higher data rate. Every OFDM symbol encodes binary data into a complex number by means of I/Q samples, as explained in Section 2.2. For instance, 16-QAM has the capacity to express 4 bits of information using just one complex number [14].

2.4 Channel State Information

CSI is the metric used in OFDM systems for describing amplitude and phase variations across subcarrier frequencies as wireless signals are transmitted between a transmitter and a receiver. To detect these variations across subcarriers, OFDM systems transmit a set of known shared pilot symbols (either interleaved in the transmitted message or as a preamble) which are then used to estimate the CSI vector H as:

$$y = H \odot x + \eta$$

where y is a vector indicating the signal detected at the receiver, x is the signal vector that was transmitted based on the agreed upon pilot symbols, \odot is the element-wise multiplication and finally η is an additive white Gaussian noise vector. H is a complex vector containing a complex number (I and Q components) for each subcarrier representing the Channel Frequency Response (CFR), from which we can retrieve the amplitude and phase of the specific subcarrier signal. Throughout this work, two ESP32 microcontrollers are employed as TX and RX antennas [15], that perform OFDM communication with a bandwidth of 20MHz using a total of 64 subcarriers. Multiple CSI samples are collected over a time window of size W , which can be represented as the $W \times S$ matrix

$$H = \begin{bmatrix} IQ_{11} & IQ_{12} & \dots & IQ_{1S} \\ IQ_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ IQ_{W1} & IQ_{W2} & \dots & IQ_{WS} \end{bmatrix}$$

where S is the total number of subcarriers. In our case, 64 are employed, of which 12 are guard (null) subcarriers, resulting in having a final set of 52 usable informative subcarriers, at a rate of transmission of 100 packets (observations) per second.

Chapter 3

Data preprocessing

Bibliography

- [1] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, “Widar3.0: Zero-effort cross-domain gesture recognition with wi-fi,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8671–8688, 2022.
- [2] D. Yi, Z. Lei, S. Liao, and S. Z. Li, “Learning face representation from scratch,” 2014.
- [3] N. McLaughlin, J. Martinez del Rincon, and P. Miller, “Recurrent convolutional network for video-based person re-identification,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1325–1334, 2016.
- [4] P. Nithyakani, A. Shanthini, and G. Ponsam, “Human gait recognition using deep convolutional neural network,” in *2019 3rd International Conference on Computing and Communications Technologies (ICCT)*, pp. 208–211, 2019.
- [5] G. Zhang, P. Zhang, J. Qi, and H. Lu, “Hat: Hierarchical aggregation transformers for person re-identification,” 2021.
- [6] S. Sadowski and P. Spachos, “Rssi-based indoor localization with the internet of things,” *IEEE Access*, vol. 6, pp. 30149–30161, 2018.
- [7] B. Yang, L. Guo, R. Guo, M. Zhao, and T. Zhao, “A novel trilateration algorithm for rssi-based indoor localization,” *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8164–8172, 2020.
- [8] Q. Dong and W. Dargie, “Evaluation of the reliability of rssi for indoor localization,” in *2012 International Conference on Wireless Communications in Underground and Confined Areas*, pp. 1–6, 2012.
- [9] Z. Yang, Z. Zhou, and Y. Liu, “From rssi to csi: Indoor localization via channel response,” *ACM Comput. Surv.*, vol. 46, dec 2013.
- [10] D. Miklavcic, N. Pavselj, and F. Hart, “Electric properties of tissues,” *Wiley Encyclopedia of Biomedical Engineering*, vol. 6, 04 2006.
- [11] Y. Wu and W. Zou, “Orthogonal frequency division multiplexing: a multi-carrier modulation scheme,” *IEEE Transactions on Consumer Electronics*, vol. 41, no. 3, pp. 392–399, 1995.
- [12] S. M. Hernandez and E. Bulut, “WiFi Sensing on the Edge: Signal Processing Techniques and Challenges for Real-World Systems,” *IEEE Communications Surveys and Tutorials*, 2022.

- [13] N. Vljajic, “Analog transmission of digital data: Ask, fsk, psk, qam.” York University CSE 3213, Fall 2010.
- [14] L. E. Frenzel, “Chapter 8 - cell phones: It is now possible to do anything wirelessly: Talk, text, email, web browse, games, whatever,” in *Electronics Explained (Second Edition)* (L. E. Frenzel, ed.), pp. 195–216, Newnes, second edition ed., 2018.
- [15] S. M. Hernandez and E. Bulut, “Lightweight and Standalone IoT Based WiFi Sensing for Active Repositioning and Mobility,” in *21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM) (WoWMoM 2020)*, (Cork, Ireland), June 2020.