12-2016

Indoor Localization of Mobile Devices Based on Wi-Fi Signals Via Convex Optimization and Bregman Divergence

Osamah Ali Abdullah
Western Michigan University, osama.ece@gmail.com

INDOOR LOCALIZATION OF MOBILE DEVICES BASED ON WI-FI SIGNALS VIA CONVEX OPTIMIZATION AND BREGMAN DIVERGENCE

by

Osamah Ali Abdullah

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Electrical and Computer Engineering
Western Michigan University
December 2016

Dissertation Committee:

Dr. Ikhlas Abdel-Qader, Ph.D., Chair
Dr. Bradley J. Bazuin, Ph.D., Co-Chair
Dr. Abiola Akanmu, Ph.D.

INDOOR LOCALIZATION OF MOBILE DEVICES BASED ON WI-FI SIGNALS VIA CONVEX OPTIMIZATION AND BREGMAN DIVERGENCE

Osamah Ali Abdullah, Ph.D.

Western Michigan University, 2016

Indoor positioning systems (IPS) have been the subject of intense academic and industrial research due to the significance of such systems in a wide range of applications. Applications of IPS include indoor navigation and its associated user services, especially by users in large complex buildings, by emergency healthcare services to locate a patient, and people with vision impairments. IPS can also play an important role in other applications that require tracking and observation, such as those used in care for the elderly or security purposes. Therefore, much research has been focused on IPS methods that fingerprint the Received Signal Strength (RSS) of the wireless local area network (WLAN) in indoor environments, the outcome of which has resulted in a positioning accuracy of close to a 1 meter.

This dissertation presents a framework based on fingerprinting maps for indoor positioning systems along with the implementation and testing results. for each technique that was investigated under this framework. This work focuses on Bregman divergences, which are a generalized form of the well-known Kullback-Leibler divergence, suited for convex functions. Since the square root of averaging KL divergence (Jensen-Shannon divergence) is a metric parameter, a framework that

incorporates the probabilistic neural network (PNN) with Jensen-Shannon Divergence (JSD) is proposed. Based on this framework, I also investigated the Jensen-Bregman Divergence (JBD). JBD is induced by a strictly convex function generator that unifies the celebrated information-theoretic Jensen-Shannon divergence with the squared Euclidean and Mahalanobis distances. JBD is used to calculate distance by focusing on dissimilarity between classes for a reliable and accurate IPS. The proposed system was implemented and simulated in the College of Engineering and Applied Sciences at Western Michigan University. To compare and allow for validation of the proposed framework, implementation and simulations of the multivariate Kullback-Leibler divergence (KL_{MVG}) under the Probability Neural Network (PNN) scheme and under the *k*-Nearest Neighbors (*k*-NN) technique were performed.

ACKNOWLEDGMENTS

First and foremost, I thank God for giving me the power, the patience, the health and the opportunity to accomplish my work and earn my Ph.D. Degree.

I would like to thank my thesis advisor, Professor Ikhlas Abdel-Qader for her guidance to me throughout my Ph.D. study at Western Michigan University. I thank her for her encouragement and support during my research. I have gained so much knowledge and experience under her supervision that allowed me to enhance my research skills and prepared me for my professional career. I would like also to thank Dr. Bradley Bazuin, for his valuable advice and feedback that highly improved my work and thesis quality. I would also like to thank Dr. Abiola Akanmu for her insight and feedback.

I would like to extend my warm thanks for my beloved family. Especially my father since he passed away before seeing me reach this point, he was the major reason for my success in my studies and in my life. Also, I want show my infinite gratitude to my mother for her support during my study at WMU and my whole life. My thanks also go to my brother Ahmad for his support and for my beloved siblings as they always were a lovely source of support and encouragement.

Last but not least, I would like to thanks the Higher Committee for Education Development in Iraq (HCED) for their support, financially and for being my mentor. I am indebted to my country for helping me to finish my Ph.D. degree.

I would like to dedicate this work to the memory of my beloved father, and to my loving mother and siblings.

Osamah Ali Abdullah

TABLE OF CONTENTS

Table of Contents—Continued

CHAPTER

Table of Contents—Continued

CHAPTER

Table of Contents—Continued

CHAPTER

Table of Contents—Continued

CHAPTER

LIST OF TABLES

LIST OF FIGURES

List of Figures—Continued

CHAPTER I

INTRODUCTION

1.1 Background and General Overview

The positioning system has undeniably become an interesting topic of market research, especially since global positioning system (GPS) chips have been integrated into smartphones. Currently, location-based services (LBS) are not limited to just location-tracking services but also offer other services, such as shopping advice, tracking of family and friends and tourist services. Many markets in Asia and the United States have integrated the positioning service into smartphones with new types of LBS coming to the forefront of smartphone technology. Indoor positioning systems (IPS) have emerged as a hot topic of research for many different applications. Figure 1 shows the diverse categories of LBS [1].

The digital map is considered to be one of the most fundamental and essential smartphone applications. It provides many services such as tracking and routing, weather information and traffic flow in the user's area, and can also be used to recruit emergency health care providers, police officers and firemen. A lot of time and resources have been preserved by using LBS to save lives. In [2], it was stated that each year, an estimated 240 million people call "911" in the United States. The Federal Communication Commission stated that approximately 70% of these emergency calls were made from wireless phones [1, 3], whereas in Europe, about

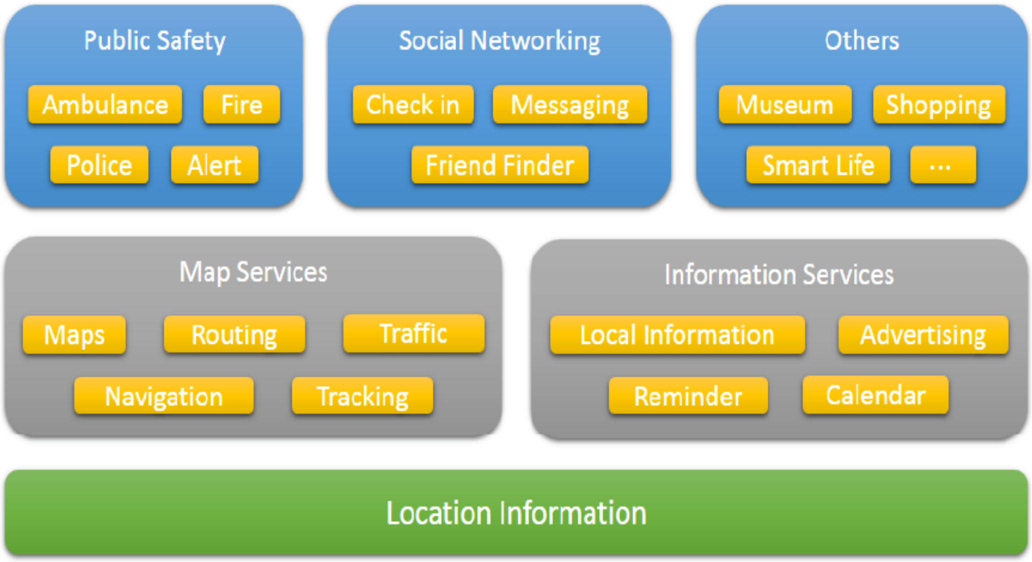


Figure 1: Categories of location-based services [1]

65% of the approximately 320 million emergency calls were wireless calls [4]. In most cases, users were unable to determine their location and guide emergency services accordingly. An additional complication arose if users had to report their location from inside a building, where phone service tends to more problematic and is less effective than phones with an outdoor positioning system. As a result, a more efficient localization technique that can provide more accurate indoor positioning coordinates is needed. LBS also have a number of uses in social networking and have become one of the biggest LBS markets with regard to the number of users and revenue, especially since the development of smartphone applications like Find My Friends and those that allow users to check in to show friends their location; these types of services have grown enormously. The main challenge that has hindered the development and application of LBS is the limitation of obtaining an accurate IPS. In [6], it was stated that people spend 80–90% of their time indoors and make about 70% of their calls indoors. Thus, in addition there is a great need to improve the performance of IPS, and, in time and as IPS improves, there will be even more applications using it.

1.2 Motivation

Estimating the coordinates of a device inside a building is a hot topic in the field of ubiquitous computing (UbiComp). Mark Weiser is considered to be the pioneer of using location information in many real-life applications, especially the ability to locate people inside buildings. While the Olivetti Research Laboratory in Cambridge, England [7], showed the possibility of building IPS, Mark Weiser was the person who saw the opportunity to use this technology for many applications. Due to the rapid growth of the UbiComp field, location computing has become a topic at the frontier of research, and estimating coordinates has become one of the most significant applications [8]. Another property of mobile devices that originated from UbiComp is context awareness, a style of computing that encompasses more than just the user's location [10], but also takes into account user identity, time and activity as the primary context types [11]. Thus, location and user activity can play important roles in UbiComp applications such as location awareness and activity recognition [1]. Even though Mark Weiser's vision has become a reality, UbiComp applications are still not very popular because of their high cost. In [12], the authors investigated this issue and found thousands of papers on Google scholar related to indoor positioning describing systems that have been proposed and implemented. After evaluating IPS, they concluded that the main challenge in deploying these systems is the cost, which is determined by three factors. First, the system has to be built in such a way to make it compatible with different devices that have different characteristics and different operating systems. Second, collecting the data to create the data map and floor plans is time-consuming, and these data are not always available. Third, in order to obtain accurate results, intensive work is required to collect data to train the locator algorithm. Even though commercial systems like Ubisense and Ekahau can achieve reasonable accuracy, they are very expensive and require customize equipment installation and maintenance. In general, there is a broad range of IPS that use different technologies, some focusing on existing ubiquitous cellular telephone or WiFi infrastructure and others that require hardware installation such as radio-frequency identification (RFID) [13], ultrasound [14], ultrawide bandwith (UWB) [5] and infrared systems [7]; systems that require costly infrastructure installation and hardware maintenance.

The GPS cannot be reliably used inside buildings for tracking and navigation mainly because the satellite signal power levels were defined based on line-of-sight receiving with minimal multipath or attenuation. In most buildings, there is no line-of-sight to the satellites and significant attenuation resulting in intermittent or total loss of the GPS signals required for position computation. Thus, an alternate source of signals is required to build an IPS, namely, new hardware that will need to be installed to obtain a new signal source such as RFID, UWD, WiFi, or cellular telephone that is readily available, easily installed and hopefully supports other useful services. The wireless local area network (WLAN) has been considered a quasi-standard for the last decade since it supports almost all handheld devices such as smartphones, computers, notebooks and tablets, all of which access the network by default. For years WiFi signals have been the main signal source in most proposed algorithms in IPS, and have been used in many algorithms for IPS to estimate the coordinates of an object with sufficient accuracy for most LBS [15]. Although these signals have some inherent complications, such as reflection, diffraction and partial absorption, the effects of these complications can be utilized by recording location-dependent signals, which are used to create a fingerprinting database called reference points (RPs). In general, location fingerprints have many advantages, but to obtain accurate results, it is mandatory that the proposed algorithm has as many RPs as possible. There are two phases of location-based fingerprinting: the offline phase and the online phase.

1.3 Challenges of Indoor Positioning

A lot of issues and challenges arise when IPS are discussed. In the past few years, many survey papers have been published that address the issues and different characteristics of these systems. Most of these papers have focused on the taxonomy or classification of LBS systems [16-19]. In the following section, the prominent prevailing challenges of IPS will be summarized. Specifically, we will focus on the issues related to location fingerprinting and will provide a broad overview of the current challenges. The related work will be covered in more detail in Chapter 2.

1.3.1 Performance

Different attributes have to be considered when evaluating IPS including the type of system and its main purpose. For example, if the system was built to track a fast-moving object, then first and foremost, it must be responsive (i.e., the time delay of the calculating position should be short). In addition, the position of the system should be optimized, as it should accurately and consistently report location coordinates [18]. As a general concern, precision and accuracy are the two main parameters that determine the performance of a system and are typically used to evaluate an IPS. Accuracy means the average error distance and precision means the probability of the cumulative distribution function with respect to position estimated. For example, in general, regular GPS receivers can determine location with an average error distance of 5 meters, indicative of accuracy, whereas precision indicates the quality of the system with approximately 90% confidence. In general, there is a trade-off between the cost and performance of IPS; the main challenge is to create a system with sufficient accuracy and reasonable cost. Higher performance means higher cost, for example, the accuracy of infrared-based positioning can be improved by adding extra filters to reduce the effects of fluorescent light [16], which will increase the price of the system.

Although fingerprinting-based location has good accuracy for most human applications such as “How do I get to Ali’s office?” the big issue remains how to use it to determine a precise locations that can be associated with a collection of fingerprinting data. The user can be disappointed by the performance of such a system, if the data collector does not share the user’s perception of the place desired.

1.3.2 Cost

Cost is one of the most important factors with regard to IPS [24], especially concerning systems that use fingerprinting location due to the time that is needed to build and maintain the radio map, in addition to other factors that require time and money such as installation and administration of the system, the maintains of a battery that the devices use, the complexity and size of the space necessary for the hardware installation infrastructure and personnel costs [17, 18].

The GPS is one of the most well-known positioning systems. Although it relies on a very large and expensive infrastructure that is very difficult and expensive to install and to maintain it is free to the users. Cost assessments of IPS are crucial and should include the factors over the lifetime of the system. Lukas, et al. [12] suggest breaking down cost into three phases: cost of installation, cost of maintenance and cost of use. As explained, hardware installation represents the primary factors that make up the majority of the cost and is determined by the choice of the hardware installation. Nevertheless, the cost will be higher if the system requires special hardware such as tags [20], antennas [21] or even WLAN access points with special capabilities [22], which makes the acquisition of data more expensive.

Another parameter that should be taken into account is provision of maps for the indoor environment, which, unlike the outdoor system, has been provided for many decades. The cost will depend upon the chosen technology. For instance, obtaining one's position using GPS technology usually requires 10 seconds and a lot of energy [12], as the user has to wait for the system to obtain the coordinates, which reduces the lifetime of the battery. On other hand, IPS use fingerprinting-based techniques and Wi-Fi radio signals. Obtaining measurements that will be used to create radio maps is a time- and energy-intensive process. Furthermore, even if the data have been scanned, the network interface may not be taken into account when the data are transferred (i.e., the user could face additional costly problems and may have to choose between having an accurate system or transferring data without considering the environment). In fact, the synchronous use of network systems is one of the biggest issues of using location fingerprinting. In addition, the total maintenance cost is difficult to assess for many systems that have been used for IPS because they have not been in service for a very long time.

In general, for systems that use the fingerprinting-based technique, the major cost is the time and the hardware needed to keep the radio map database up to date. As will be explained later, the signal can suffer from variation that affects the system’s results; thus, a system update is inevitable. The cost will be high if updates of the training data are performed by special personnel or must be done manually. Another factor that increases the cost of the fingerprint-based technique is the cost of the installation and maintenance of APs. Furthermore, the received signal strength (RSS), expressed in dBm, perceived can differ from device to device, which are backed up by hardware network adapters. For example, the signal strength of a mobile phone can be 47 dBm, while that of a laptop is 65 dBm at the same location. This will be elaborated upon more in the next section.

1.3.3 Signal Variation

The IPS was designed and built to serve specific applications, for instance, tracking objects that are moving, as well as the usual user demands for high accuracy. Some systems require installation of specific, expensive hardware (e.g., UWB) for accurate indoor positioning to within 10 centimeters. On the other hand, WLAN signals do not need installation because almost every building currently has Wi-Fi, but these radio signals are subject to multipath, refraction, humidity, temperature, reflection and many other factors. As a result, the RSS varies and fluctuates over time and has a low correlation with distance, which leads to inaccurate coordinate measurements [18].

Furthermore, if the device that records the signal is moving, it can sense spatial variations on both a small and large scale. In general, the strength of a signal depends on how it propagates; for example, the farther one moves from the source, the weaker the signal becomes, resulting in the inaccurate estimation of coordinates. Systems that use fingerprinting-based location tend to have better results, because it will reduce the effect signal variations, including those related to wall absorption and reflection. People can cause temporal variation in signals, albeit small-scale fluctuations, because the human body is a good antenna that can act like a sink for signals. Another factor that can cause signal variation is the weather, which depends on the climate zone, as well as the change of seasons. All of these factors are challenges that require management of the associated cause and effects. In summary, fingerprinting-based localization can make use of any signal, but in order to guarantee a certain level of accuracy over time, it is mandatory that the training data remain updated, which raises the question of how this can be accomplished without significantly increasing the cost.

1.4 Goals and Contributions

There are two main problems with IPS: 1) RSS fluctuations vary with time for many reasons including reflection, absorption, multipath, temperature, and the motion or presence of people; and 2) It is difficult to generate a radio map with rich characterization in the offline phase. The more measurements that are taken, the better the accuracy and the easier it is to obtain data with high precision, but this work is labor-intensive and time consuming [1]. A primary goal of this work is to enhance the accuracy of fingerprinting-based localization systems by first and foremost understanding the main problems that have persisted with IPS. To this end, the first objective will be to analyze and understand the cause and effects of signal variation. This will allow training data to be generated with high accuracy location that are cost effective, easy to access, and can be maintained and used for a long time. For this reason, different collection methods have been used and tested in different algorithms. The main purpose of this work was to use collaborative fingerprinting that could be easily updated and would not require specialized personnel, and to use an algorithm that could deal with the multimodal distribution of signals to measure the similarity between test and training data. A procedure with high characterization of size distribution was proposed to solve the first problem. Specifically, RSS values were taken in four different orientations (45°, 135°, 225° and 315°) to prevent body-blocking effects, with 10 scans and a delay of 10 seconds to reduce the effects of signal variation.

The chapters of this Dissertation are organized as follows. After presenting our introdcution in Chapter 1, the main concepts of IPS have been explained in addition to the common terminology that will be used as the basis to evaluate and present the positioning system. Throughout Chapter 2, related, and an overview of the technology used in the IPS system will be provided. In work will be analyzed in detail chapter two, the concept and different methodologies used in location fingerprinting will be explained in depth. In Chapter 3, Bregman divergence will be analyzed in detail. I developed a powerful and flexible method for constructing metrics from convex functions in indoor positioning system. In chpater 4, we propose a Probabilistic Neural Network (PNN)-based method as a tool for matching fingerprinting in the active phase with Jensen-Shannon Divergence (JSD) as the measure of similarity. By constructing the probability kernels that we used in a weighted regression scheme, we will be able to match user location to a fingerprint. In Chapter 5, we propose a Probabilistic Neural Network (PNN)-based method as a tool for matching fingerprinting in the active phase with Jensen-Bregman Divergence (JBD) as the measure of similarity. In Chapter 6 we propose a meta-algorithm that combines k-mean clustering with JSD to find the similarity and estimate the object's location. In Chapter 7, Jensen-Bregman Divergence (JBD) is proposed for a WLAN-based method. We perform the matching stage using probability kernels as a regression scheme and compare the results with JSD and Kullback-Leibler multivariate Gaussian distribution to measure the similarity between the RSS measurement of test points and RPs to investigate RSS behaviours. Finally, chapter 8 summarizes this work and future work.

CHAPTER II

EXISTING SYSTEM AND ALGORITHM IMPLMENTATION: LITERATURE REVIEW

2.1 Background and General Overview

Prior to development of positioning technology, different signals were used to obtain one’s physical location. However, the development of positioning systems has led to different technologies that have been a great success for estimating outdoor location such as the global positioning system (GPS), which uses radio frequency (RF) tracking technology. Other technologies such as gyroscopes and accelerometers, which are considered precise inertial measurement devices, have been used in airplanes and missiles for accurate navigation and localization. Currently, wireless signals including light and sound waves and RF signals can be used to estimate location in indoor environments. Microelectromechanical sensors (MEMS), including inertial and pressure sensors, have been integrated into smartphones. The driving force behind developing these technologies has been to make constant advancements in wireless communication technologies. Due to the different types of technologies and measurements, different kinds of techniques have been used to estimate the location of an object; the major location estimation algorithms are proximity, fingerprinting, triangulation, and dead reckoning. Other possible measurement quantities, such as angle-of-arrival (AOA), involve time-of-flight (TOF), link quality, and sensor readings, although many challenges in positioning systems still need to be solved in relation to these technologies. In different situations, these algorithms have unique advantages and disadvantages, which led us to propose a framework that incorporates them and also provides better performance. The main goal for all researchers is to obtain better performance with high accuracy and low cost, although other parameters such as coverage, robustness, and complexity also need to be taken into account. In this chapter, the literature on IPS will be reviewed to provide an overview of the positioning systems used in this proposal. First, components of IPS will be described. Then different techniques and technologies will be briefly described, and finally, location-based IPS will be discussed.

2.2 Common Localization Systems

The technologies used for indoor localization can be divided into four categories: RF, acoustic waves, light waves, and MEMS sensors.

2.2.1 RF technologies

Various RF technologies are widely used in indoor positioning systems such as augmented GPS, RFID tag, Wireless Personal Area Networks (WPANs), Bluetooth, WSNs, FM radio signals, UWB, and Wi-Fi.

2.2.1.1 Augmented GPS

GPS is considered to be one of the most globalized localization systems. However, although it provides an accurate and reliable positioning service for outdoor applications [51], it does not work in indoor environments. As a result, multiple approaches have been proposed, such as providing a new hardware design, to address this shortcoming and improve GPS performance to make it feasible for indoor positioning [52–55]. Although these experiments have improved GPS accuracy to within 6 to 16 m [56], it is still inefficient in indoor environments, and in some environments, the GPS signal is completely blocked. In [57], the authors proposed a local station and antenna that can repeat the GPS signal, known as GPS pseudolites. Transmitting using ground based pseudo-satellite transmitters within the GPS system can be achieved using carrier-phase or code-phase signals with high accuracy as small as 1cm. The indoor GPS receiver can acquire the same levels of pseudo-range and code-phase signals as the outdoor receiver, and localization to within a few centimeters can be achieved with little error. In [40], the authors proposed to use GPS repeaters and modified delay lock loop GPS receivers for sequential switching, such that the error in distance localization is reduced to a few centimeters. Even though GPS pseudolites and repeaters are considered to be accurate indoor localization systems, they require a precise time synchronization, and are susceptible to multipath interference.

2.2.1.2 Cellular Networks

In general, cellular networks are distributed over land areas called cells, each served by at least one fixed-location transceiver, known as a cell site or base station. Cellular technology is classified into different mobile telecommunication technologies such as 2G, 3G, 4G, and long-term evolution (LTE). The cellular network has good signal penetration that can be used for indoor localization. Furthermore, it has wide coverage, long-term operational stability, and an existing infrastructure. The RSS is a measurable quantity for Global System for Mobile Communications (GSM) signals, which be used for both outdoor and indoor localization systems. In [58], a correlation database was used in an urban outdoor environment to compare the RSS measurements of the serving cell with six neighboring cells. The authors reported a localization distance error of 80 m for the 67th percentile and of 192 m for the 95th percentile. In

[60], the kNN was used to estimate the localization of an object by matching the online GSM fingerprints with the offline fingerprints; each fingerprint consisted of 35 channels with accuracy ranging from 1.94 to 4.07 m. In code division multiple access (CDMA), networks such as Universal Mobile Telecommunications System (UMTS) and 3G control the transmitted power to accommodate the network load, which can increase signal variation and lead to increases in localization distance errors. In [59], a CDMA indoor localization system was proposed that applied a fingerprinting technique named CILoS based on signal delay instead of RSS. CILoS captures the relative time difference at which signals emanate from different base stations at specific locations because the CDMA system needs precise synchronization to provide distance localization accuracy within 5 m. Furthermore, the UMTS network has a poor indoor penetration signal that can be affected by multipath propagation and signal attenuation, resulting in IPS that can only provide accuracies within hundreds of meters [61–63]. The major problem with using 3G and 4G technologies for indoor localization systems is that the network uses high-frequency bands that can cause multipath propagation signals.

2.2.1.3 UWB

UWB technology has a very low transmitted radio signal. The large bandwidth of the UWB allows it to support wireless data rate from 480 Mbps up to 1.6 Gbps, over distances of several meters. The advantage of UWB technology over other technologies used for indoor positioning is that it requires very little power. Because UWB has a narrow band, UWB pulses have the ability to effectively penetrate walls, which makes them less susceptible to multipath propagation [65]. The location technology that UWB is based upon is very efficient with low localization distance error because it allows precise time and propagation measurements using UWB pulses. In [64], the authors proposed enhanced TDOA (E-TDOA) by considering a differential impulse response (DIR) in the TDOA domain for indoor localization, with a localization distance error of less than 1 m. It was reported in [66] that the accuracy of RSS-based lateration ranges from 0.1 to 0.2 m, but the disadvantage is that it is expensive compared with other technologies.

2.2.1.4 RFID and NFC

RFID systems are generally composed of RFID tags that store user information and a reader that consists of a transceiver that records the data from the tags, each of which is modulated by adding a unique identification code. There are two kinds of tags: passive and active. The passive tags take energy from the receiving signal, and the active tags are battery-operated. There are four different RFID frequency bands: low frequency (LF) (125 kHz), high frequency (HF) (13.56 MHz), ultra-high frequency (UHF) (433, 868–915 MHz), and microwave frequency (2.45 GHz, 5.8 GHz). In general, RFID tags are used for indoor localization and play an important role in the “the deployment of an indoor positioning scheme. In [67], the authors used RFID technology and adopted a weighted proximity algorithm to localize the moving RFID reader. In [68], an AOA algorithm was proposed with a passive RFID and robust phase characterization to estimate the angle. Near field communication (NFC) is a special branch of RFID technology that has recently drawn attention because it is used in smartphones. NFC is a more secure way to exchange data over short distances (5 cm or less), and has been used in indoor localization systems when tags are placed in areas of interest [69].

Furthermore, NFC supports peer-to-peer communication, and is currently used in places such as museums and stores.

2.2.1.5 Bluetooth

Bluetooth is a WPAN technology that is used in indoor localization systems, and works at 2.4 GHz. It has a wide range of applications especially with portable devices since it performs multiple functions by interacting with other devices. Bluetooth devices transmit signals that occupy 1 MHz of spectrum and use a frequency-hopping, spread-spectrum mechanism to spread signals. Because Bluetooth communication hops among channels, narrowband interference that only blocks a few channels does not have significant overall impact on communications [70]. Nevertheless, this technology is susceptible to frequency interference from other devices that work on the same frequency band. Most Wi-Fi indoor localization algorithms have been applied to Bluetooth technology for indoor positioning purposes [34, 35, 47]. In [34], multiple neural networks were used with Bluetooth technology by connecting the smartphone to multiple Bluetooth systems. A hybrid framework was proposed by Subhan *et al.* [47], who combined RSS-based lateration with RSS fingerprinting-based localization that depended on Bluetooth technology. The log-distance path loss (LDPL) parameters were estimated depending upon the fingerprinting results, whereas lateration methods were used to estimate the coordinates of the object. In general, Bluetooth has low transmitter power so many Bluetooth tags are needed to cover a small area. The accuracy of this technology rivals that of Wi-Fi technologies [35], although the biggest challenge with Bluetooth technology is its latency, as it needs at least 10 seconds to process inquiries [36]. In addition, its use causes a noticeable decrease in battery life. Recently, another version of Bluetooth, called Bluetooth low energy (BTLE), was developed for indoor localization purposes and has a battery life that exceeds one year. Furthermore, it needs less time to process inquiries (milliseconds) and can transmit signals over a longer distance, which makes it a good choice for multiple applications, such as indoor localization. iBeacon was produced by Apple for indoor positioning purposes, and is used in iPhone with BTLE technology. This technology is based on the proximity technique, which estimates the location coordinates using the range of the device and proximity algorithm, instead of using longitude and latitude. The iBeacon is not restricted to IPhones, and can also be used with android systems. For example, the TI sensor tag can support both android smartphones and Apple iBeacon [37]. In addition, the NOKIA Corporation research center produced a new technology for indoor location based on Bluetooth technology using an AOA algorithm called High Accuracy Indoor Positioning. It is a modified version of BTLE technology with an antenna that can yield very precise accuracy with a localization distance error of around 0.5 m.

2.2.1.6 Wi-Fi

Wi-Fi is a technology that represents the WPAN protocol based on the 802.11 IEEE network standard. It operates with two frequency bands, namely, 2.4 and 5 GHz Industrial, Scientific, and Medical (ISM) radio bands, and is considered to be one of the most popular systems used for wireless data communication. It can be used at work, at home, in the airport, and in fact in almost any building, and has a data rate of up to 1300 Mbit/s with a range of up to 70 m or more. IPS that use Wi-Fi technology are being extensively studied. Due to the fact that Wi-Fi does not need an infrastructure, it does not entail any extra costs. Thus, it is an appealing technology, particularly because it provides more coverage than RFID and Bluetooth. However, the AoA, ToA, TDoA are less commonly used for indoor localization due to multipath signal interference, and the difficulty of obtaining time synchronization and angular measurements [38]. Most IPS that use Wi-Fi technology have incorporated the fingerprinting-based technique, as detailed in section 2.3.

2.3 Computational Methods for Localization

RSS-based IPS techniques have been classified into two major categories: lateration and fingerprinting-based techniques. In general, the lateration method often leads to inaccurate estimations. It was reported in [71] that the typical office is 200 ft. long and 80 ft. wide, with an average localization error of about 24.73 ft. This may be because of the mulitpath propagation model (PM) due to the non-line-of-sight challenge and sensitivity to errors in even one access point (AP) coordinate estimate. The lateration method uses log-distance to describe the relationship between mobile devices and multiple APs [72]. The accuracy of lateration-based localization is very sensitive to multiparameters, for example, if one AP coordinate is inaccurate, the performance of this technique can decrease dramatically.

Figure 2 shows a functional block diagram of an IPS, consisting of a number of location-signal sensing devices, a positioning algorithm requiring spatial signal information, and a display system. First, the signal is detected by sensing devices at an unknown location using different technologies such as infrared, RF, or ultrasound.

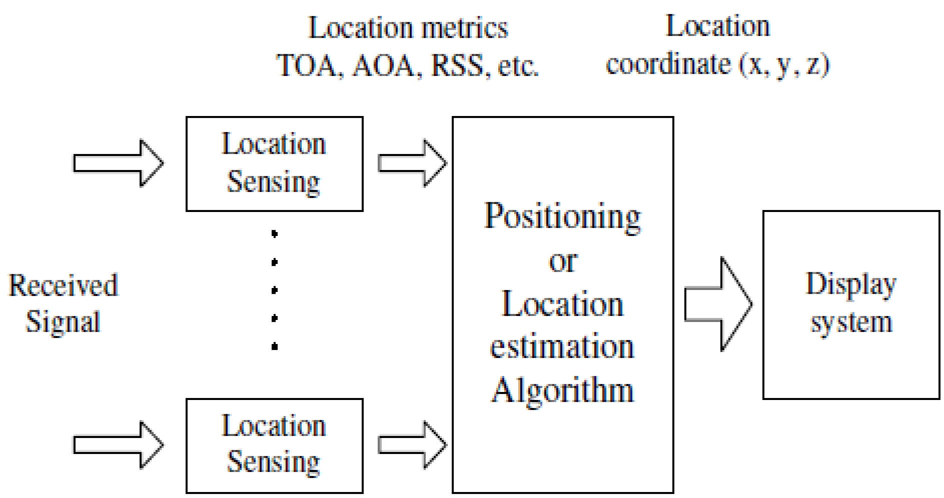


Figure 2: A functional block diagram of the IPS [26]

Then the signal is converted into location metrics depending upon the available characteristics of the sensed signal such as angle of arrival (AOA), time of arrival (TOA), received signal strength (RSS), or carrier signal phase of arrival (POA) [25]. Multiple sensor signal characteristics may also be collected to form unique patterns at locations. Next, the location is estimated using different algorithm approaches such as the distance-based approach [27], signal processing [28], probabilistic approach [29], or neural networks [30]. Finally, the location is displayed on the display system. Hightower *et al.* [31] proposed that the location system could be considered a location stack framework, which is a software engineering model that divides the positioning problem into smaller research problems.

2.3.1 Triangulation

Triangulation is a technique used to estimate the coordinates of an object based on the geometric properties of a triangle. It has two implementations: lateration and angulation. In general, the lateration technique estimates the coordinates of an object by measuring the distance of the target from multiple reference points (RPs) and using different scalar quantities such as time-of-flight and RSS. The number of algorithms depends on time, such as time difference of arrival (TDOA), TOA, and round-trip time of flight (RTOF). The angulation technique estimates the coordinates of an object by measuring the angle of the target with respect to multiple RPs, also known as the AOA.

2.3.1.1 TOA

The TOA technique estimates the distance between two points by measuring the propagation time between them. There should be at least three TOA measurements from three different RPs. The coordinates are estimated using the intersection point of the RPs in the circle of radiation as illustrated in Figure 3; however, shadowing and multipath will occur, which can lead to inaccurate distance estimation.

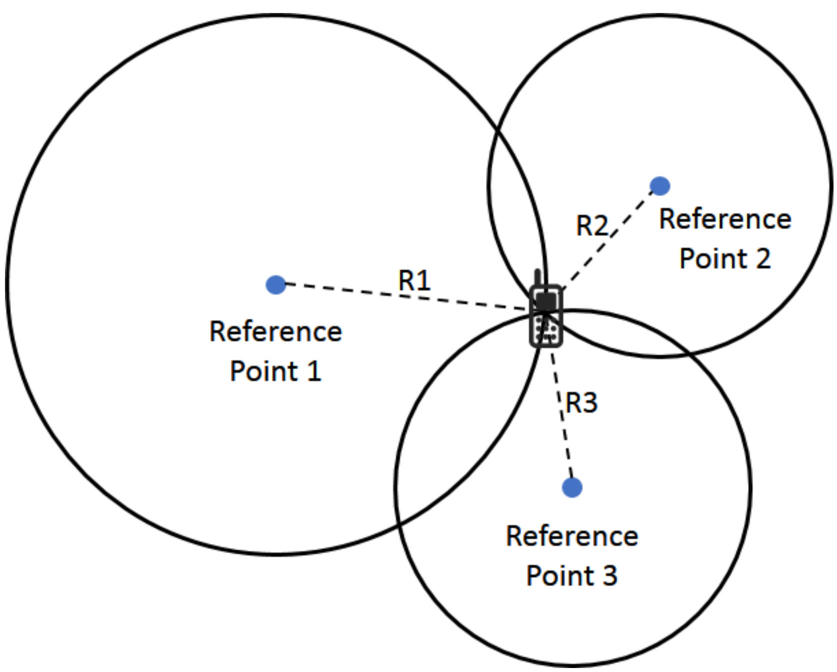


Figure 3: Localization based on TOA measurements

Maximum Likelihood (ML) and least squares (LS) estimation methods are the most common statistical procedures used to decrease these estimation errors and estimate the target coordinates. If we let (x_m,y_m) indicate the unknown location of mobile coordinates, and (x_1,y_1), (x_2,y_2), and (x_3,y_3) indicate the known reference coordinates, the distance between reference i and the target will be calculated by:

$$r_i = (t_i - t_0)c \tag{2.1}$$

where c represents the speed of light, and t_0 and t_i represent the transmission and reception time of the signal, respectively. After applying the geometric relationship, the following sets of equations will be obtained:

$$(x_1 - x)^2 + (y_1 - y)^2 = R_1^2 \tag{2.2}$$

$$(x_2 - x)^2 + (y_2 - y)^2 = R_2^2 \tag{2.3}$$

$$(x_3 - x)^2 + (y_3 - y)^2 = R_3^2 \tag{2.4}$$

where R1, R2, and R3 represent the radius of the radiation circle. This algorithm can be improved by taking three RPs and considered as 3D [32]. The LS method can be used to solve this overdetermined system of nonlinear equations. If we subtract (2.2) from (2.3) and (2.4), the formula will become:

$$r_2^2 - r_1^1 = x_2^2 - 2x_2x_m + y_2^2 - 2y_2y_m - x_1^2 + 2x_1x_m - y_1^2 + 2y_1y_m \tag{2.5}$$

$$r_3^2 - r_1^1 = x_3^2 - 2x_3x_m + y_3^2 - 2y_3y_m - x_1^2 + 2x_1x_m - y_1^2 + 2y_1y_m \tag{2.6}$$

Equations (2.5) and (2.6) can be rewritten as:

$$\begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \end{bmatrix} = \frac{1}{2} \begin{bmatrix} K_2^2 - K_1^2 - r_2^2 + r_1^2 \\ K_3^2 - K_1^2 - r_3^2 + r_1^2 \end{bmatrix} \tag{2.7}$$

$$K_1^2 = x_1^2 + y_1^2 \tag{2.8}$$

Equation (2.7) can thus be written as

$$Hx = b \tag{2.9}$$

where

$$H = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{bmatrix}$$

$$x = \begin{bmatrix} x_m \\ y_m \end{bmatrix}$$

$$b = \frac{1}{2}\begin{bmatrix} K_2^2 - K_1^2 - r_2^2 + r_1^2 \\ K_3^2 - K_1^2 - r_3^2 + r_1^2 \end{bmatrix}$$

Then, the LS method is applied for the final estimation using the following formula:

$$\hat{x} = (H^T H)^{-1} H^T b \tag{2.10}$$

Alternatively, the ML method can be applied to TOA measurements by finding the probability that distances r_1, r_2, and r_3 find the target location x as follows:

$$\hat{x} = \arg\max\{P(r_1, r_2, r_3 \mid x)\} \tag{2.11}$$

Since distances r_1, r_2, and r_3 are independent, the joint probability will become:

$$P(r_1, r_2, r_3 \mid x) = P(r_1 \mid x)P(r_2 \mid x)P(r_3 \mid x)$$

Estimation of the distance between the RPs and target are considered a Gaussian error distribution, and consequently the ML can be rewritten as:

$$\begin{aligned} \hat{x} &= \left(\frac{1}{\sqrt{2\pi\sigma_1}} e^{\frac{-(d_1-r_1)^2}{2\sigma_1{}^2}} \frac{1}{\sqrt{2\pi\sigma_2}} e^{\frac{-(d_2-r_2)^2}{2\sigma_2{}^2}} \frac{1}{\sqrt{2\pi\sigma_3}} e^{\frac{-(d_3-r_3)^2}{2\sigma_3{}^2}} \right) \\ &= \arg\min \sum_{i=1}^{3} \frac{-(d_i - r_i)^2}{2\sigma_i{}^2} \end{aligned} \tag{2.12}$$

Time synchronization in the TOA technique is very important as it can provide high accuracy if the time from the transmitter and receiver is synchronized correctly. Unfortunately, the multipath signal has a large effect on time synchronization, which can lead to a system with low accuracy. In general, the TOA and related TDOA technique are widely used including ultra-wideband (UWB) [27–31], where it has good penetration that reduces the effects of the multipath signal.

2.3.1.2 TDOA

The TDOA technique estimates the coordinates of the target by measuring the difference in TOA from multiple RPs. Specifically, it measures the difference in distance between the RPs and the mobile target based on the time difference as shown in Figure 4. From a geometric point of view, TDOA measurements are provided for mobile targets that lie on a hyperboloid with a constant difference between two RPs. TDOA is used to determine the position of a target by sending two different kinds of signals and to find the difference in TOA at multiple measuring units in order to assign the transmitting node's position; the equation is:

$$\frac{R}{C_1} - \frac{R}{C_2} = t_1 - t_2 \tag{2.13}$$

where C_1 represents the speed of the first signal, C_2 represents the speed of the second signal, R is the distance from the transmission unit to the target, and t_1 and t_2 are the times that it takes for these signals to reach the target [32].

One advantage of TDOA over TOA measurements is that only RPs need to be synchronized. As example applications, TDOA is used in GPS systems [39, 40], UWB [42, 43], and wireless signal networks (WSNs) [41].

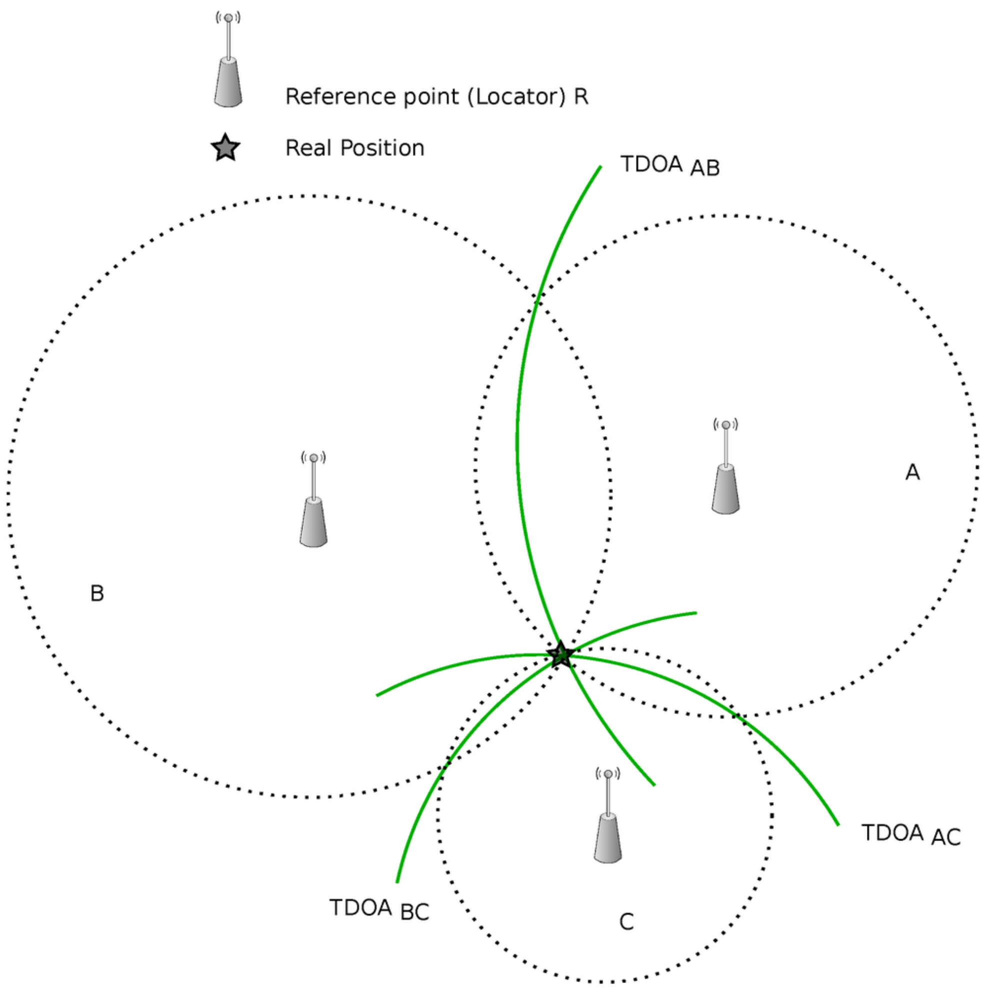


Figure 4: Localization based on TDOA measurements

2.3.1.3 RSS-based Lateration

The RSS-based technique estimates the coordinates of an object by calculating the attenuation signal between the RPs and the mobile target. In general, due to the multipath signal, it is very difficult to obtain time synchronization between the target and the RPs. However, RSS-based lateration is relatively standard and can increase the efficiency of IPS. It is also very popular in many indoor positioning technologies such as Wi-Fi [44–45], UWB [66], radio frequency identification (RFID) [48], and Bluetooth

[47]. The RSS-based technique estimates distance by estimating the signal path loss due to propagation. In an indoor environment, shadowing and multipath are very complex to solve and can directly affect the performance of the system. Thus, for long-range signals, RSS-based lateration is not a suitable approach, although its performance is much better for short-range signals. Once the distance is estimated, the coordinates can be easily estimated following the same procedure as that used to calculate the TOA.

2.3.1.4 AOA

AOA is an angulation technique used to estimate the coordinates of a mobile object by interpreting the intersection point of pairs of hypothetical signal paths along particular angles, as shown in Figure 5. In general, AOA only needs two RPs to localize the mobile object, but for three-dimensional (3D) localization, three RPs are needed. Its accuracy can be improved if three or more RPs are used in a technique referred to as multi-angulation. With the AOA technique, antenna array or directional antenna are used to estimate the lines-of-bearing for the system [40, 41]. The advantage of AOA over the above mentioned techniques is that only two RPs are needed and time synchronization is not needed between the receiver and transmitter. On the other hand, the AOA system needs complex antenna elements and processing hardware to obtain the coordinates measurement, and it is difficult to obtain precise angle measurements in an indoor environment due to shadowing and multipath signals.

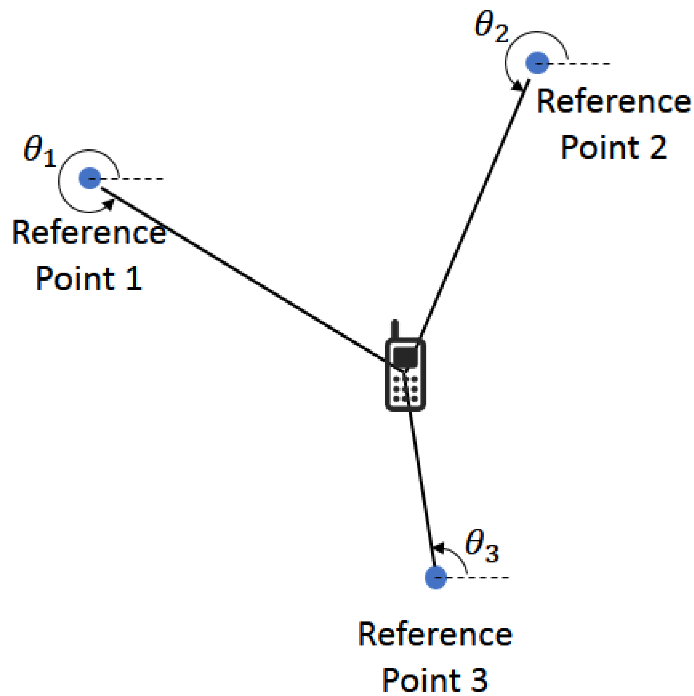


Figure 5: Localization based on AOA measurements

2.4 Fingerprinting-based Localization

Fingerprinting "discovered" based on applying estimation theory and pattern matching concepts and algorithms to RSS lateration data and systems with significantly enhanced performance. Advantages of such system included less dependence on accurate RSS mean power estimates allowing better performance in non-line-of –sight and multipath environments.

Filter techniques have been used to ensure localization performance, for example, in [73] the particle filter was proposed, in [74] the Kalman filter was used, and in [75] the Bayesian filter was used to ensure performance by restricting localization errors through trace movements. Thus, instead of solving the PM problem to estimate the location of the object, a radio map was prebuilt to use in the fingerprinting-based localization [76] scheme, which consists of two phases: an offline and an online phase. The offline phase collects RSS readings with time sampling and location to generate a prior fingerprint at the fingerprinting database, which contains the RPs. The number of RPs directly impacts the performance of fingerprinting-based methods. The online phase estimates the actual location by using IPS to compare the RSS value of a mobile device with predefined fingerprints. One of the simplest ways to estimate a mobile user's location is the k-nearest neighbor's algorithm (kNN), which estimates localization by computing the k-nearest neighbors that have the smallest Euclidean distance between the two phases [77], as shown in Figure 6. Such an algorithm has low accuracy but is easy to implement. The fingerprinting-based technique is based on the difference in RSS between the mobile device and the RP fingerprints [78]. A modified probability neural network was used in [36] with kNN vector mapping to estimate the location of the object based on the RSS values of the RPs and the reported results of the object. Recently, a probabilistic approach was proposed in [79] and [80] by developing a RSS-based probabilistic technique and estimating the location using kernel methods. In [79], the probability was estimated using a Bayesian framework, whereas in [80] a histogram framework was used. In [81], a Kullback-Leibler (KL) divergence framework was used with a composite hypothesis to formulate the localization problem. In [82], KL divergence was used to estimate the probability density of the RSS of the unknown device. In addition, the RSS from APs was treated as a multivariate Gaussian distribution in that study. It was reported in [83] that the probability of KL divergence as a kernel regression scheme can achieve error margins of up to 1 m in office environments, but the authors used Bluetooth technology for non-Gaussian fingerprints.

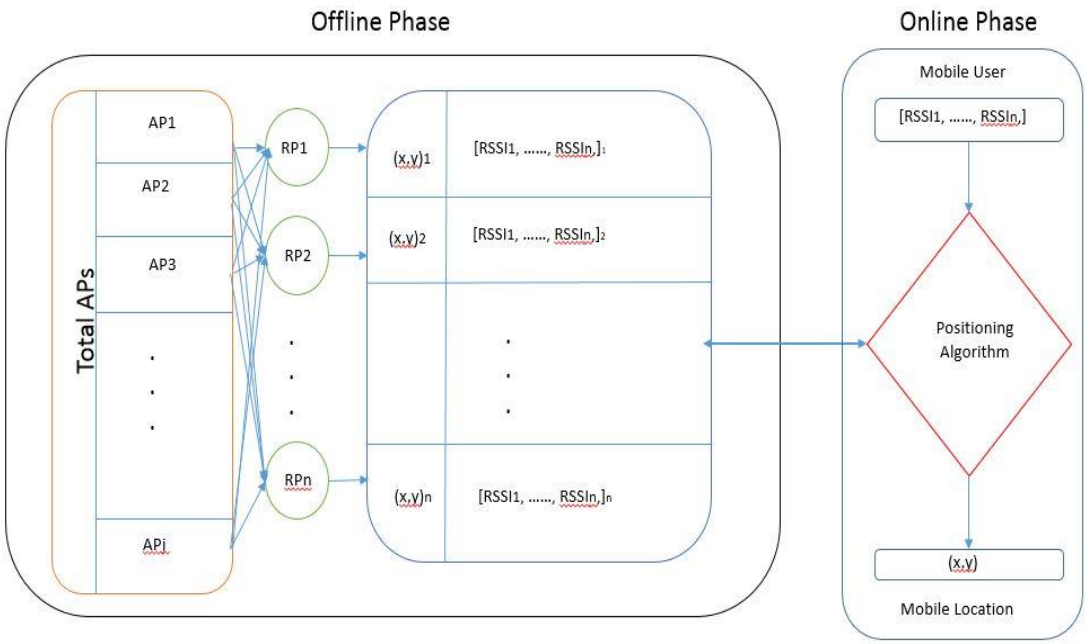


Figure 6: Architecture of the proposed hierarchical IPS

CHAPTER III

BREGMAN DIVERGENCE FRAMEWORK

3.1 Introduction

Bregman divergences are based on convex optimization, recently they received great attention, they ae well known as the generalization of the Kullback-Leibler divergence. Bregman divergence can introduce a class of “squared root metrics”, which can be regarded as natural generalization of Euclidean distance. For a variety of machine learning applications, the interests have been gone beyond Euclidean distance in the recent years. Bregman divergence are a multi-class of distortion functions, such as, Kullback-Leibler divergence, Euclidean distance, Mahalanobis distance, Itakura-Saito distance, etc. Bregman divergences are not symmetric and that considers as crucial property. Thus, the algorithm that exploit scalability properties lay beyond the using of Bregman divergence. In the recent year, a lot of attempts have been done to investigate the symmetry of Bregman divergence and satisfy the triangle inequality that will lead to squared metrics.

3.2 Bregman Formulation

Recently, the study of measuring the distortion in classes is increasing instead of depending on a single distance. This trend is witnessed in many applications in machine learning, computational geometry and IPS. Measuring the similarity/dissimilarity using the Bregman divergence became attractive these days because it encapsulates both the information-theoretic relative entropy and the geometric Euclidean distance, which is a meta-algorithm [93]. The Bregman distance D_{φ} between two sets of convex space data p = (p_1, . . . , p_d) and q = (q_1, . . . , q_d) that is associated with φ (which is defined as a strictly convex and differentiable function) can be defined as:

$$D_{\varphi}(p,q) = \varphi(p) - \varphi(q) - \langle \nabla \varphi(p), p - q \rangle \tag{3.1}$$

where $\langle .,. \rangle$ denotes the dot product and

$$\langle p,q \rangle = \sum_{i=1}^{d} p^{(i)} q^{(i)} = p^{T} q \tag{3.2}$$

and $\nabla\varphi(p)$ denotes the gradient decent operator:

$$\nabla \varphi(p) = \left[\frac{\partial \varphi}{\partial p_1} \cdots \frac{\partial \varphi}{\partial p_d} \right]^{T} \tag{3.3}$$

The Bregman divergence unifies the statistical Kullback-Leibler divergence with the geometry squared Euclidean distance by defining the distortion measurement in classes:

- ❖ The Euclidean distance was obtained from the Bregman divergence by considering the convex function as $\varphi(p) = \sum_{i-1}^{d} p_i^{2} = \langle p, p \rangle$, which is the parabolic potential function in figure 7.
- ❖ The Kullback-Leibler divergence is also a Bregman divergence if the convex function used is $\varphi(p) = \sum_{i-1}^{d} p_i \log p_i$ which is defined as negative Shannon

entropy. The Kullback-Leibler divergence is defined for two discrete distributions as:

$$KL(p||q)=\sum_s p(S=s)\log(\frac{p(S=s)}{q(S=s)}) \tag{3.4}$$

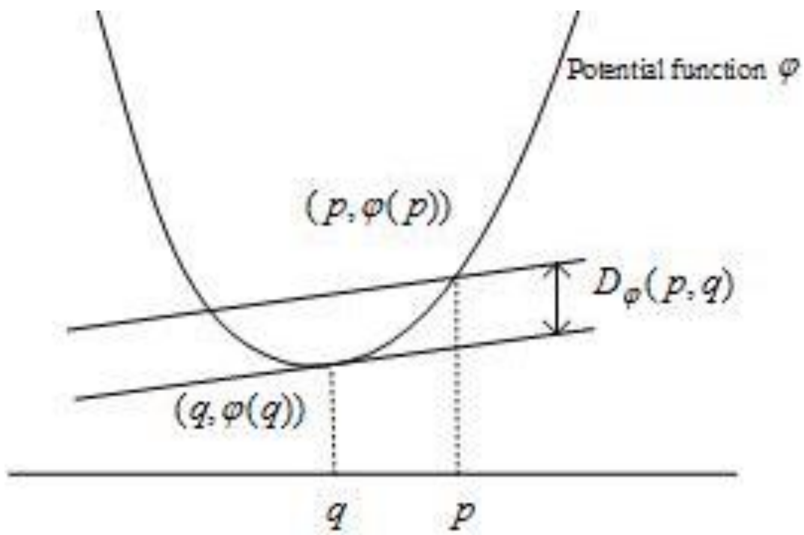


Figure 7: The Bregman divergence represents the vertical distance between the potential function φ and the hyperplane at *q*

In information theory, the Shannon differential entropy measures the amount of uncertainty of a random variable:

$$H(p) = p\log\frac{1}{p} \tag{3.5}$$

The Kullback-Leibler divergence is equal to the cross-entropy of two discrete distributions minus the Shannon differential entropy [24]:

$$KL(p||q) = \sum_s H^x(p(S=s) \parallel (q(S=s) - H((p(S=s) \qquad (3.6)$$

where H^x is the cross-entropy:

$$H^x(p(S=s) \parallel (q(S=s) = \sum_s p(S=s)\log\frac{1}{q(S=s)} \qquad (3.7)$$

where S is the set of vectors of the RSS. In general, the Bregman divergence is not symmetric, but it can symmetrize as follows:

$$SD_\varphi(p,q) = \frac{D_\varphi(p,q) + D_\varphi(q,p)}{2} \qquad (3.8)$$

$$= \frac{1}{2}\langle p-q, \nabla\varphi(p) - \nabla\varphi(q)\rangle \qquad (3.9)$$

In the same manner, Jeffreys divergence symmetrizes the oriented Kullback-Liebler divergence as follows:

$$J(p,q) = KL(p \parallel q) + KL(q \parallel p) \qquad (3.10)$$

$$= H(p \parallel q) + H(q \parallel p) - (H(p) + H(q) \qquad (3.11)$$

$$= \sum_s ((p(S=s) - q(S=s))\log(\frac{p(S=s)}{q(S=s)} \qquad (3.12)$$

Such an information-theoretic divergence has two major drawbacks: first, the output can be undefined if q=0 and p≠0, and secondly, the J-divergence is not bounded in terms of metric distance, so to avoid these drawbacks, and avoid the log(0) or divide by 0, [115] propose a new divergence called the K-divergence :

$$K(p \parallel q) = KL(p, \frac{p+q}{2}) \tag{3.13}$$

By introducing the K-divergence, [93] produce the Jensen-Shannon divergence (JSD) by depending on the K-divergence as:

$$JSD(p \parallel q) = \frac{1}{2}(KL(p, \frac{p+q}{2}) + KL(q, \frac{p+q}{2})) \tag{3.14}$$

$$= \frac{1}{2}(H(p \parallel \frac{p+q}{2}) - H(p) + H(q \parallel \frac{p+q}{2}) - H(q)) \tag{3.15}$$

$$= \frac{1}{2}(\sum_{i=1}^{L} p_i \log \frac{p_i}{\frac{1}{2}\overset{(\circ)}{q}_{j,i} + \frac{1}{2}p_i} + \sum_{i=1}^{L} \overset{(\circ)}{q}_{j,i} \log \frac{\overset{(\circ)}{q}_{j,i}}{\frac{1}{2}\overset{(\circ)}{q}_{j,i} + \frac{1}{2}p_i}) \tag{3.16}$$

The JSD can be always (1) defined, (2) bounded by an L1-metric, and (3) finite. In the same vein the Bregman divergence can be symmetrized as:

$$SD_{\varphi}(p, q) = \frac{1}{2}(D_{\varphi}(p, \frac{q+p}{2}) + D_{\varphi}(q, \frac{q+p}{2})) \tag{3.17}$$

$$= \frac{\varphi(p) + \varphi(\overset{(\circ)}{q}_j)}{2} - \varphi(\frac{p + \overset{(\circ)}{q}_j}{2}) \tag{3.18}$$

for d-dimensional multivariate data:

$$SD(p, q) = \sum_{i=1}^{L} \frac{\varphi(p_i) + \varphi(\overset{(\circ)}{q}_{j,i})}{2} - \varphi(\frac{p_i + \overset{(\circ)}{q}_{j,i}}{2}) \tag{3.19}$$

where q represents the fingerprint data set, and p, the dataset of the test points, represents the APs the mobile received. Since φ is a strictly convex function and the $SD(p,q)$ equal to zero if and only if p=q , this family of distortions is termed Jensen-Shannon divergence. The geometric interpretation is represented in figure 8, where the divergence represents the vertical distance between $((\frac{p+q}{2}), \varphi(\frac{p+q}{2}))$ and the midpoint of the segment $[(p, \varphi(p)), (q, \varphi(q))]$.

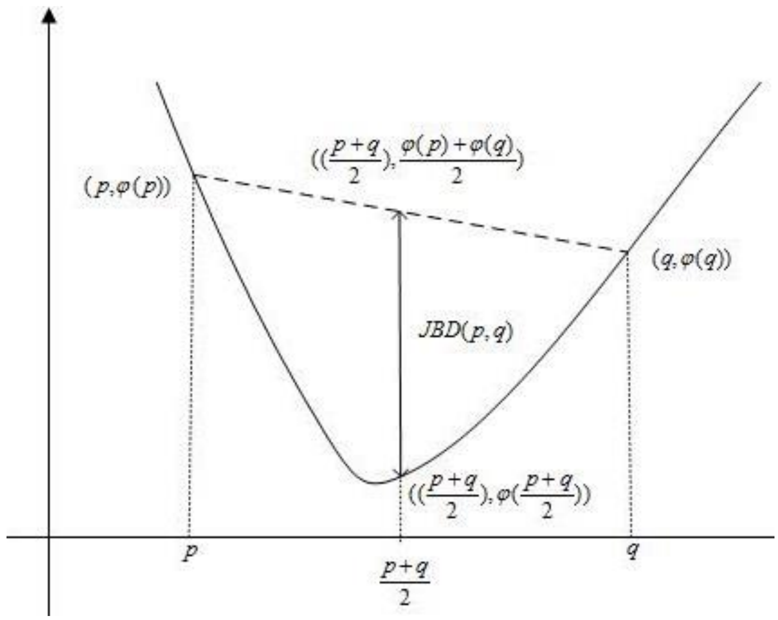


Figure 8: Interpreting the Jensen-Bregman divergence.

In general, for a positive definite matrix $Q \succ 0$, the Jensen-Bregman divergence has all the quadratic distance $\varphi(p) = \langle Qp, p \rangle$, which is known as Mahalanobis distance.

$$
\begin{aligned}
SD(p,q) &= \frac{\varphi(p)+\varphi(q)}{2} - \varphi(\frac{p+q}{2}) \\
&= \frac{2\langle Qp, p\rangle + 2\langle Qq, q\rangle - 2\langle Q(p+q), p+q\rangle}{4} \\
&= \frac{1}{4}(\langle Qp, p\rangle + \langle Qq, q\rangle - 2\langle Qp, q\rangle) \\
&= \frac{1}{4}\langle Q(p-q), p-q\rangle \\
&= \frac{1}{4}\|p-q\|_Q^2
\end{aligned}
$$

Table 1 contains a list of some convex optimization functions with their corresponding Bregman divergences. Bregman divergences have many interesting properties that can be useful in many machine learning applications, such as non-negativity, convexity in the first argument, etc.

Table 1: Bregman divergences generated from some convex functions.

Domain	$\varphi(p)$	$D_{\varphi}(p,q)$	Divergence
$\Re$	p^2	$(p-q)^2$	Squared loss
$\Re_+$	$p\log p$	$p\log(\frac{p}{q})-(p-q)$	
[0,1]	$p\log p+(1-p)\log(1-p)$	$p\log\frac{p}{q}+(1-p)\log(\frac{1-p}{1-q})$	Logistic loss
$\Re_{++}$	$-\log p$	$\frac{p}{q}-\log(\frac{p}{q})-1$	Itakura-Saito distance
$\Re$	e^p	$e^p-e^q-(p-q)e^p$	
$\Re^d$	$\lVert p\rVert^2$	$\lVert p-q\rVert^2$	Euclidean distance
$\Re^d$	$p^T Ap$	$(p-q)^T A(p-q)$	Mahalanobis distance
d-Simplex	$\sum_{j=1}^{d} p_j \log_2 p_j$	$\sum_{j=1}^{d} p_j \log_2(\frac{p_j}{q_j})$	KL-divergence
$\Re_+^d$	$\sum_{j=1}^{d} p_j \log p_j$	$\sum_{j=1}^{d} p_j \log_2(\frac{p_j}{q_j})-\sum_{j=1}^{d}(p_j-q_j)$	Generalized I-divergence

CHAPTER IV

A PROBABILITY NEURAL NETWORK-JENSEN-SHANNON DIVERGENCE FOR A FINGERPRINT BASED LOCALIZATION

4.1 Introduction

The main usage of indoor positioning systems is to estimate the location of persons and objects, which is very useful for many applications such as in logistics, health, and construction industries to name a few. Indoor localization based on received signal strength (RSS) has been also used for multiple asset management. Researchers have proposed different indoor localization technologies and reported different localization accuracies. These include systems that employ technologies such as ultrasonic, Bluetooth, ultra-wideband, radio-frequency IDs, and RSS [85]. Among these numerous technologies, due to its availability, RSS in a wireless local access network (WLAN) has become a research focus in the recent years.

Reported RSS-based indoor localization techniques can be categorized into two classes: fingerprinting and lateration based techniques. However, the lateration method is suffering from the inaccurate estimation of the location of the object in indoor environments.

As an example, in [86] it was reported that the average localization error is about 24.73 ft in a typical office scenario with a width of 80 ft and a length of 200 ft. Such a performance is suffering due to two main reasons: 1) the log distance propagation model (PM) used, which is a consequence for the non-line of sight (NLOS) problem [74] intrinsic in lateration techniques, and 2) lateration methods are sensitive to errors when one or more AP positions are not accurately estimated. As a result, in recent years, fingerprinting-based localization has become dominant in IPS research [78]. The profile of a given indoor environment and the RSS of some predetermined reference points (RPs) with its coordinates are all acquired and used to generate fingerprints in the prior or offline phase. Thus, in the fingerprint database, each fingerprint is associated with a physical location in the building.

The number of RP per unit area is obviously considered one of the main parameters that impact fingerprinting based methods performance. Fingerprinting based techniques are based on collecting RSS measurements that are very labor-intensive and time-consuming. Fingerprinting is accomplished in an off-line phase. The online phase is when the RSS profile of a mobile device will be compared to predefined fingerprints to determine its location using the developed system. Other paramount parameters of performance accuracies are selecting an appropriate method to measure the distance, such as kNN.

In this chpater, we propose a Probabilistic Neural Network (PNN) based method as the tool to match fingerprinting in the active phase combined with Jensen-Shannon divergence (JSD) as the measure of similarity. By constructing the probability kernels that we used in a weighted regression scheme, the matching of user location to a fingerprint will be achieved.

4.2 Related Work

The probabilistic approach to received signal strength indicators (RSSI) has recently been proposed for IPS in [98] and [81]. They proposed to model the RSSI distribution of each location as a probabilistic framework and to solve for location estimates using

Kernel Methods. Also, a Bayesian framework has been used to estimate the probability of having a Bayesian Network [98] or a specific histogram of RSSI at a new location [81]. In [36] they used the modified probabilistic neural network with k-nearest neighborhood vector mapping for indoor location estimation based on received signal strength. In [82], the localization problem was formulated as a composite hypothesis testing where each hypothesis is associated with a family of probability density functions (PDF) based on a Kullback-Leibler (KL) divergence framework. They used KL Divergence to estimate the corresponding probability densities of the unknown position with the signature of each cell and showed that the accuracy can be improved by considering the "impact" of neighboring cells closest to the correct one and eliminating the distant or incorrect cells. In their work, they assumed the RSSI from multiple APs is simply a multivariate Gaussian distribution [83]. In [84], they used KL divergence in the probabilistic kernel regression of the location achieving up to 1 m accuracy in office environments for non-Gaussian fingerprints by using a Bluetooth localization technique. In [89] they focused on the modeling of RSSI distribution, the Gaussian and non-Gaussian distributions, by using probabilistic kernels for comparing Gaussian distributions and finding their limitations by using KL-divergence kernel regression. However, the KL divergence can easily compute the joint distributions of multiple independent multinomials. An automated adjustment method has been proposed in [90] for a signal shift from an unknown device when they pass through easy-to-estimate locations, such as doorways, but the primary drawback of this method is that one may not enter the easy-to-estimate locations.

4.3 Bregman Divergence Algorithm Formulation

The probability distribution measurement was increased because of the importance of measuring the distance in the environment that suffers from inference and discrimination. The concept of measuring distance by using two probability distributions was initially proposed by Mahalanobis [91]. Bregman distance [92] is used to measure distance by focusing on dissimilarity between classes and encapsulating the information in relative entropy and Mahalanobis distance. Bregman divergence is defined using Taylor series expansion as follows:

- Let φ: Δ → R be a real-valued strictly convex function defined on a closed convex set Δ. The Bregman divergence is defined as

$$D_{\varphi}(p,q) = \varphi(p) - \varphi(q) - \langle \nabla \varphi(p), p - q \rangle \tag{4.1}$$

where $\nabla \varphi(p) = \left[\frac{\partial \varphi}{\partial p_1} \ldots\ldots \frac{\partial \varphi}{\partial p_d} \right]^T$ denotes the gradient operator, and $\langle p, q \rangle = p^T q$ denotes the inner product [43], and D_{φ} at point p is the first order Taylor expansion of φ at point q [44]. If we choose $\varphi(p) = \sum_{i-1}^{d} p_i \log p_i$ (which is defined as the negative entropy), equation 1 leads to KL divergence. If $\varphi(p)$ has been chosen as $\sum_{i-1}^{d} p_i^2 = \langle p, p \rangle$, equation 1 will lead to Euclidian distance [93] and the generalized quadratic pseudo distance (Mahalanobis distance) will be:

$$D_{\varphi}(p \parallel q) = (p - q)^T Q(p - q) \tag{4.2}$$

where Q is a positive definite symmetric matrix and represents the variance-covariance matrix of the data set. The Bregman divergence can be interpreted as the vertical distance between the tangents at q that pass through p as shown in figure 9.

$$D_{\varphi}(p||q)=H(p,q)-H(p) \tag{4.3}$$

where $H(p)$ is the entropy of p and $H(p,q)$ is the cross-entropy due to using q instead of p. In the discrete case where variable S takes a discrete value (e.g. RSSI values from access points), we have

$$KL(p||q)=\sum_{s} p(S=s)\log(\frac{p(S=s)}{q(S=s)}) \tag{4.4}$$

where KL divergence is a non-symmetric measurement, i.e., in general $D_{\varphi}(p,q) \neq D_{\varphi}(q,p)$, and can only be symmetric for a positive definite matrix if and only if $p=q$.

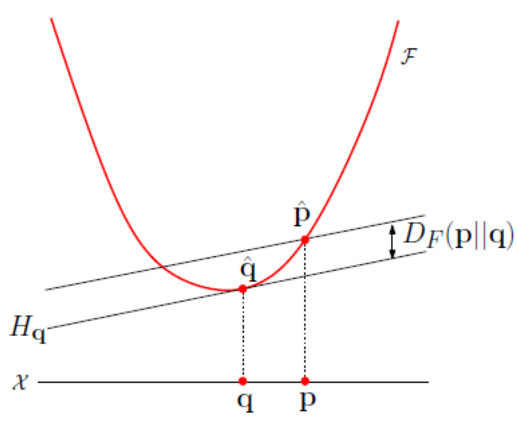


Figure 9: Geometric interpretation of Bregman divergence as the vertical distance between the tangents at q that pass through p [45].

The symmetrized KL divergence D between two distributions p and q can be defined as [13]:

$$D(p,q) = KL(p \parallel q) + KL(q \parallel p) \tag{4.5}$$

which can be undefined because the KL divergence is undefined if q=0 and p≠0, which means both values should be continuous. To overcome this issue, [91] proposed a new directed divergence method between p and q called the Jensen-Shannon divergence (JSD) to overcome these drawbacks:

$$\begin{aligned} JSD(p \parallel q) &= \sum\nolimits_s p(S=s)\log\frac{p(S=s)}{\frac{1}{2}q(S=s)+\frac{1}{2}p(S=s)} \\ &+ \sum\nolimits_s q(S=s)\log\frac{q(S=s)}{\frac{1}{2}q(S=s)+\frac{1}{2}p(S=s)} \end{aligned} \tag{4.6}$$

4.4 PNN-JSD IPS Method

PNN is a supervised feed-forward network derived from Bayesian Decision Networks. The input vector is tested and trained with the dataset data to normalize it into the network. The first layer consists of a neuron of the input features while the second layer has a neuron (pattern unit) for each feature of the training dataset. The pattern unit will be computed using dot products of the training pattern and input pattern. PNNs use radial basis function (RBF), activation function in the hidden layer (second layer), which is used to make a decision on a sample of the input [50]. In a probability distribution, a fingerprint considers the conditional probability distribution of signal strengths given the cell position indexed by ℓ is $q_\ell = q(S_j \mid \{x_\ell, y_\ell\}$. Because of the need to use multiple access points (APs), S is a multivariate random variable. The Bayes conditional probability rule is given by:

$$P(\{x_i, y_i\}/S_j) = \frac{q_\ell P(\{x_i, y_i\})}{\sum_{j=1}^{m} P(S_j/\{xj, y_j\}) P(\{x_j, y_j\})} \quad (4.7)$$

where $P(\{x_i, y_i\}/S_j)$ is the conditional probability density function (PDF) of S_j given a set of multivariate RSSI samples acquired at fingerprint cell i and $P(\{x_i, y_i\})$ is the probability of drawing data from the class $\{x_i, y_i\}$.The input vector S_j (candidate for classification) is assumed to belong to class $\{x_i, y_i\}$ if

$$P(\{x_i, y_i\}/S_j) > P(\{x_j, y_j\}/S_j) \qquad \forall j = 1,2,3,....m \quad (4.8)$$

The challenge with the prior probability $P(S_j/\{xj, y_j\})$ is that it is unknown. To address this issue, a probability kernel-based approach has been used [38].

$$P(p, q) = \frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{(p-q)^2}{2\sigma^2}) \quad (4.9)$$

where σ is a nonzero number that can be chosen by the user. There is no theoretical basis for choosing an optimal σ. The Kernel function will be equal to 1 if $p = q$ and start to decay if the dissimilarity between these two inputs becomes larger. The architecture of Probabilistic Neural Network Indoor Positioning System is shown in figure 10 with layers.

4.4.1 Input Layer

This layer represents the test point for which an estimate of its location is sought. The number of neurons in this layer is equal to the number of APs (variables) that is needed to describe the form of the input.

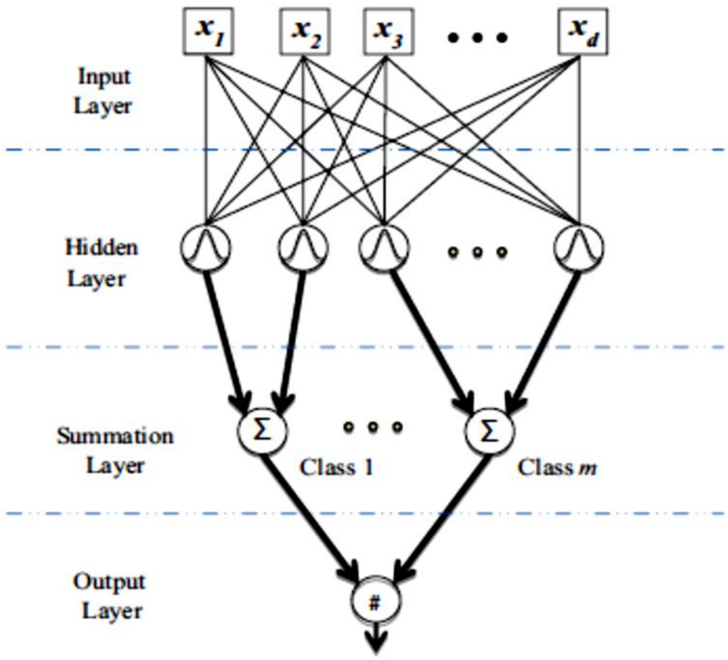


Figure 10: PNN Architecture adapted in this work and as presented in [10].

4.4.2 Pattern (hidden) Layer

The learning set has been organized in the pattern layer by representing each input vector with a hidden neuron that records the parameters of the current input vector. PNNs use the training dataset to estimate PDFs using Reference Points (RPs) and then used to estimate the likelihood that the input sample belongs to a given class as shown in algorithm 1 [50]. In this case, we can make an assumption of local independence of different AP distributions at location {x,y}

$$p(S \mid \{x,y\}) = \prod_{j=1}^{J} p(S_j \{x,y\}) \qquad (4.10)$$

Algorithm 1. The PNN positioning method

1- During offline phase, collect the RSSI measurement from APs at specific locations to generate the fingerprinting map.

2- During online phase, collect the RSSI from the APs at the unknown position, set the APs in the same way as the database of the offline phase with respect to the similar MAC address.

3- During online phase, perform the following steps for each fingerprint cell:

   - Estimate the smoothing factor σ that maximizes the output.
   - Estimate the similarity/dissimilarity using Eq. 3.9.

4- Transfer the maximum outputs to the Output Layer.

4.4.2.1 Modified Training Algorithm

A probability kernel-based approach using the symmetrized JSD method is proposed. The probability kernels can be implemented using either a regression scheme or simply a weighted scheme that allows for estimates of these probabilities from the training samples and lead to estimates of the location of the object. The JSD divergence is used to measure the distance between the estimated probability distribution and the true probability distribution [97], then is used to estimate the likelihood of the input sample as belonging to a given class. The JSD of a joint distribution of independent variables is equal to the sum of the JSD for the signal distribution according to the chain rule of relative entropy. In this case, we can make an assumption of the local independence of different AP distributions at location {x,y}. The RSSI distribution at fingerprints can be defined as

$$kJSD(p,q_{\ell}) = \exp\left(-\frac{JSD(p(S_j\{x,y\}), q(S_j\{x_{\ell},y_{\ell}\}))}{2\sigma^2}\right) \tag{4.11}$$

To improve the accuracy, we propose the PNN-JSD positioning method as presented in Algorithm 2:

Algorithm 2. The PNN-JSD positioning method

1- During offline phase, collect the RSSI measurement from APs at specific locations to generate the fingerprinting map.
2- During online phase, collect the RSSI from the APs at the unknown position, set the APs in the same way as the database of the offline phase with respect to the similar MAC address.
3- During online phase, perform the following steps for each fingerprint cell:
    - Estimate the smoothing factor σ that maximizes the output.
    - Estimate the similarity/dissimilarity using Eq. 3.11.
4- Transfer the maximum outputs to the Output Layer.

4.4.3 Output Layer

The number of units must equal the number of the existing classes in the output layer. This layer is also called the summation layer. Each neuron in this layer is connected to all the neurons in the pattern layer. The output probability class for the output layer represents the class that the input belongs to. The outputs of summation layers are proportional to the density functions, which can be seen as probabilistic output. But to obtain more accurate results and estimate the location with more proficiency, we should consider more than one class and perform Weighted Kernel Regression [84] to obtain an estimate of the location using p, the sampled distribution of RSSI:

$$(\bar{x}, \bar{y}) = \frac{\sum_{\ell} (x_{\ell}, y_{\ell}) k(p, q_{\ell})}{\sum_{\ell} k(p, q_{\ell})} \tag{4.12}$$

From this equation, one notices that instead of using the whole set of recorded training data points, we use only K nearest neighbors. Kernels provide a way to interpret the location estimates between fingerprint locations.

4.5 Implementation and Experiment Results

This section provides details on the experimental evaluation of the proposed positioning system. The positioning software was developed in Java using Eclipse Framework version 4.2, and installed on an HP Pavilion to provide the localization service. Furthermore, a Samsung S5 mobile phone with an Android 4.4.2 operating system was used for RSS sampling during data collection. The proposed work was implemented in an academic building, namely the first floor of the College of Engineering and Applied Science building at Western Michigan University (WMU). This area includes a large study lounge and three big rooms and a long corridor in the area of 23.5 m × 16.5 m. The layout is shown in figure 11.

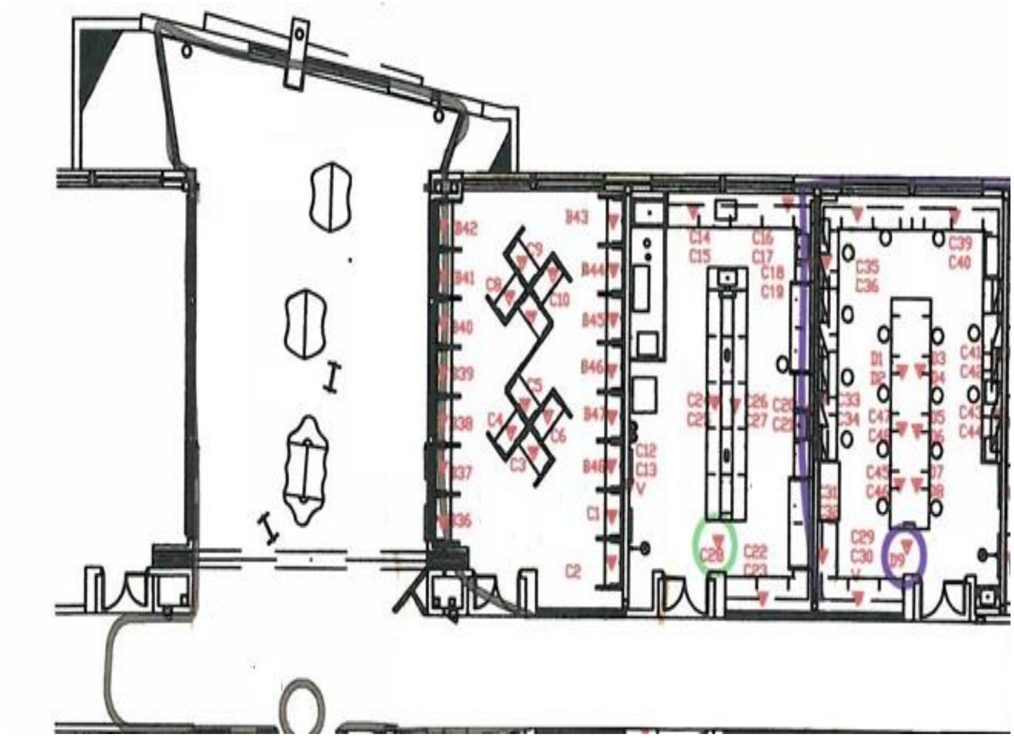


Figure 11: Tracking results on the 2D layout for the first floor of the College of Engineering and Applied Sciences at WMU.

A 2-D coordinate system is used to describe the eighty-four RPs, which have been created in a 1.5m × 1.5m grid. Cisco Linksys E2500 Advanced Simultaneous Dual-Band Wireless-N Routers were used as APs. As the phone was held by a person at a height around 4.2’, the body blocking effect and data acquisition time couldn’t be ignored. The flow of people changed with time, and we did not control the people’s movement in each office room. Generally, during the day, more people walked around and in and out of the room, which affected the distribution of the signals. We truly needed a procedure that could provide a richer characterization of the distribution. Therefore, in our experimental simulations, the recording of RSSI was taken with four different orientations (45 °, 135 °, 225 °, and 315°). Ten scans were taken in the same place with a delay of 10 seconds, and four orientations in the same place to prevent the body-blocking effects. The fingerprint map was generated by taking the average reading of four orientations and of 10 recordings at each orientation.

The number of APs that covers an area of interest can play an important role to improve the quality of positioning systems. In this section, we will investigate the impact of the variation of the number of APs on the RP neighbors that will be used to estimate the position of the object and how that can affect the accuracy of the Wi-Fi system.

To validate our proposed work performance and allow for a comparison of results obtained, we also implemented PNN and kNN methods as presented in [36] and [78]. The numbers of nearest neighbors (NN) that will be used to estimate the location of the object are 5, 20, and 80.

Figure 12 shows the localization error when using the 5 nearest neighbors for kernel regression under different AP selections. Obviously the best results were obtained when using 22 APs. The performance of the positioning system decreases when more APs are used. The highest median accuracy was 1.12 m for PNN-JSD, 1.34m for PNN and 1.38 m for kNN.

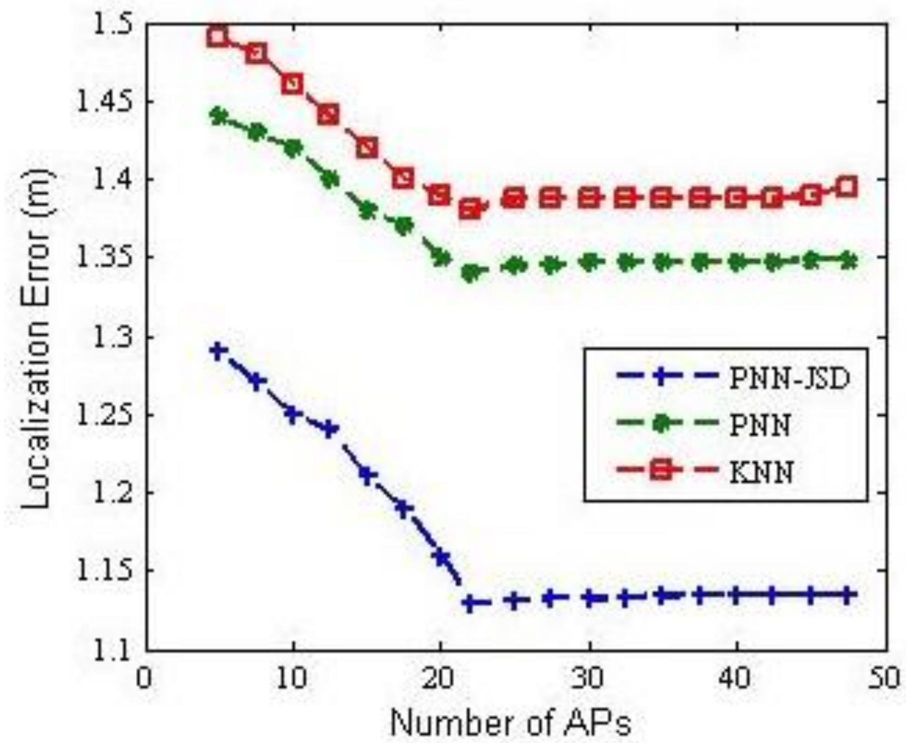


Figure 12: Localization distance error of PNN-JSD versus different algorithms with different number of APs with 5 NN

Figure 13 shows the localization error when using the 20 nearest neighbors for kernel regression under different AP selections. Least localization errors have been obtained for all the systems. The highest accuracy happened when 22 APs were used. The highest median accuracy was 1.01 m for PNN-JSD, 1.097 m for PNN and 1.19 m for kNN. A further decrease was noticed in the localization error distance when using 80 nearest neighbors for kernel regression under different AP selections as shown in figure 14. The highest median accuracy was 0.96 m for PNN-JSD, 0.995 m for PNN and 1.12 m for kNN. Table 2 shows the comparison of the proposed positioning system with the different fingerprinting approach, such as kNN [78], kernel-based method [87], and compressive sensing [88]. The proposed method showed the 90% error gave the lowest distance error between the other algorithms.

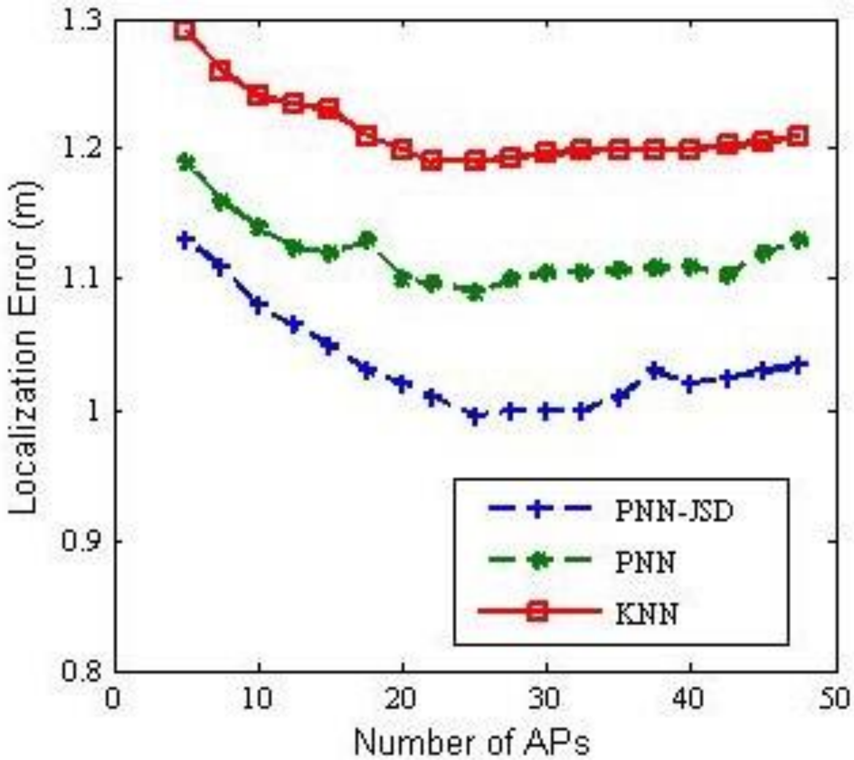


Figure 13. Localization distance error of PNN-JSD versus different algorithms with different number of APs with 20 NN

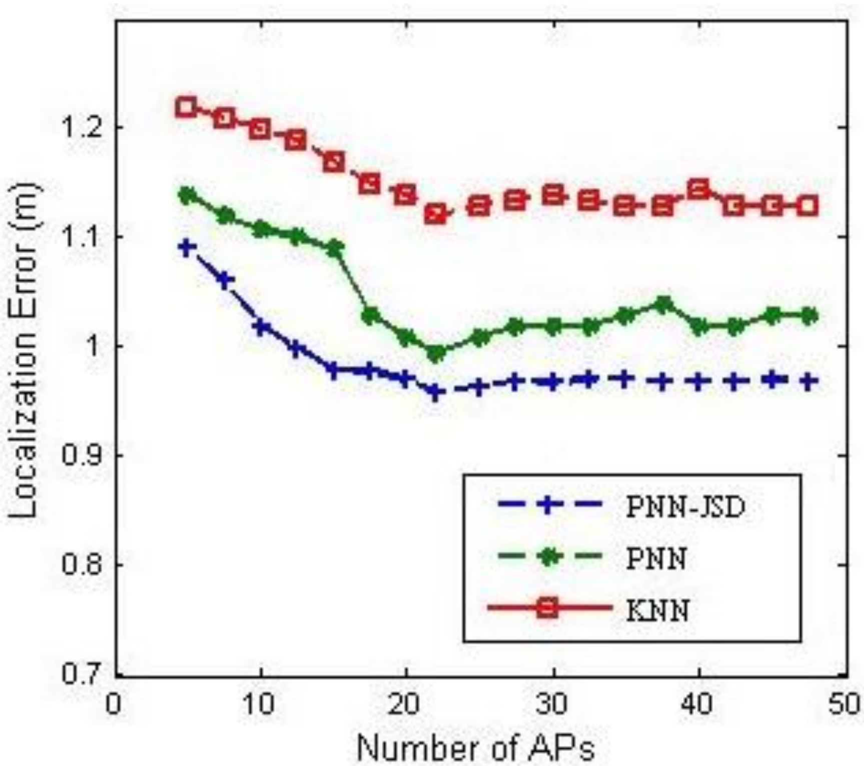


Figure 14: Localization distance error of PNN-JSD versus different algorithms with different number of APs with 80 NN

Table 2. Localization distance error of PNN-JSD versus different algorithms

Technique	Median [m]	Accuracy 90% [m]
kNN	1.8	3.7
Kernel-based	1.6	3.6
CS-based	1.5	2.7
PNN-JSD	0.98	1.93

4.6 Conclusion

Indoor localization can be used as a navigation tool in normal conditions but also in abnormal conditions such as in emergency healthcare services or while in unfamiliar buildings where people can get disoriented or lost easily. We created a fingerprint map for a segment of our college utilizing the spatial relation of RSS readings. We used the PNN scheme and integrated it with the JSD method. We also compared the results with the results of other approaches such as PNN and kNN distance. The results showed that our proposed scheme is a feasible alternative for IP systems. It has the advantage of requiring a moderate amount of effort in collecting and training data with an average error of less than 1 meter. This result is adequate for an indoor environment under normal conditions. The PNN-JSD method results have higher accuracy than PNN and both produced slightly better accuracy than the kNN stand-alone method. We also have found that taking all 80 neighbors into consideration thus far results in best estimates.

We are in the process of investigating position prediction error distributions and in need to quantify the localization variation of the Wi-Fi signal distribution in space.

CHAPTER V

CONVEX OPTIMIZATION VIA JENSEN-BREGMAN DIVERGENCE FOR WLAN INDOOR POSITIONING SYSTEM

5.1 Introduction

Demands for pervasive and mobile computing systems are increasing exponentially, in particular in smartphones. Indeed, such demands have made it practical to provide Location-Based Services (LBSs) like IPS and navigation [98]. To provide a high level of accuracy is challenging due to the complexity of indoor localization data. Different algorithms have been proposed by researchers, and different estimations of accuracy have been obtained using RSS, Bluetooth, radio-frequency IDs, ultra-wideband, or ultrasonic technologies [85]. Due to its availability and lower cost, the utilization of WiFi received signal strength (RSS) has become the focus of research in recent years.

RSS-based IPS techniques have been classified into two major techniques: Lateration and fingerprinting-based techniques. In general, the lateration method often suffers from inaccurate estimation. It was reported in [85] that within the typical office of a length of 200 ft. and a width of 80 ft., the average localization error is about 24.73 ft. This may be due to the mulitpath propagation model (PM) due to the non-line of sight challenge and the sensitivity to errors in even one AP coordinates estimate; Thus, instead of dealing with the propagation model problem to estimate the location of the object, a radio map was prebuilt to use in a fingerprinting-based localization [77] scheme. In general, the fingerprinting-based localization consists of two phases: an offline and online phase. The offline phase collects the RSS readings with their time sampling and their location to generate a prior fingerprint at the fingerprinting database, which contains the reference points (RPs). The number of RPs has a direct impact performance of fingerprinting-based methods. The online phase estimates the actual location by comparing the RSS value of a mobile device with the predefined fingerprints by using an IPS. One of the simplest ways to estimate the mobile user's location is the k nearest neighbor algorithm (kNN), which estimates the localization by computing the k nearest neighbors that have the smallest Euclidean distance between the two phases [78]. Such an algorithm has low accuracy but is easy to implement. In this work, we propose:

- A Probabilistic Neural Network-Jensen-Bregman Divergence (PNN-JBD) for a WLAN-based method. We perform the matching stage using probability kernels as a regression scheme.

- A procedure with high characterization distribution to be used. RSS value was taken in four different orientations (45 °, 135 °, 225 °, and 315°) to prevent body-blocking effects, with ten scans with a delay of 10 seconds to reduce the effect of signal variation.

- PNN-JBD results outperforms the results of PNN and kNN with respect to accuracy and the average error distance, which indicates that the proposed combining scheme is more effective in sensitive environments of WLAN-based positioning systems.

5.2 Related Work

There are two main types of IPS methods: lateration and fingerprinting. The lateration method uses log-distance to describe the relationship between mobile devices to multiple access points (APs) [86]. The accuracy of lateration-based localization is very sensitive to multi-parameters, such as if one AP coordinate has been taken inaccurately this can drop the performance dramatically. The filter techniques have been used to ensure the localization performance; for example, in [74] the particle filter was proposed, while in [99] the Kalman filter was used and in [76] the Bayesian filter was used to ensure the performance by restricting the localization error through trace movements. The fingerprinting-based technique is based on the difference in RSS between the mobile device and the RP fingerprints [78].

A modified probability neural network was used in [100] with kNN vector mapping to estimate the location of the object based on RSS of the RPs and the reported results of the object. Recently, a probabilistic approach was proposed in [80] and [81] by developing a RSS based probabilistic technique and estimating the location by using kernel methods. In [80 the probability was estimated by using a Bayesian framework while in [81] a histogram framework was used. In [82], a Kullback-Leibler (KL) divergence framework was used with a composite hypothesis to formulate the localization problem. In [83], KL divergence was used to estimate the probability density of the RSS of the unknown device. Also in their work, RSS from the APs was treated as a multivaritate Gaussian distribution.

It was reported in [84] that the probability of the KL divergence as a kernel regression scheme can achieve up to 1m error distance measurement in office environments, but they used Bluetooth technology for non-Gaussian fingerprints.

5.3 Bregman Divergence Algorithm Formulation

Achieving high accurate location estimates in indoor positioning systems using probability distribution models is found to be challenging due to the variation and interference of which RSS suffers from. Initially, measuring a distance by using a probability distribution was proposed by Mahalanobis and followed by different types of distance measures proposed by others such as in [91]. As an example, Bregman distance [92] appears attractive because it measures the distance by encapsulating the information-theoretic relative entropy and the geometric Euclidean distance. A Bregman distance D_φ measures the distortion between classes that is defined by a Jensen convexity gap that is induced by a strictly convex function φ as in equation 1:

$$D_\varphi(p,q) = \varphi(p) - \varphi(q)\langle\nabla\varphi(p), p-q\rangle \tag{5.1}$$

where $\langle .,. \rangle$ denotes the inner product and

$$\langle p,q\rangle = \sum_{i=1}^{d} p^{(i)}q^{(i)} = p^T q \tag{5.2}$$

where $\nabla\varphi(p)$ denotes the gradient operator of φ at point q:

$$\nabla\varphi(p) = \left[\frac{\partial\varphi}{\partial p_1}......\frac{\partial\varphi}{\partial p_d}\right]^T \tag{5.3}$$

The geometric representation of Bregman divergence is represented in figure 15. The Bregman divergence measures the distortion by unifying the parameters of the Kullback-Leibler divergence and squared Euclidean distance:

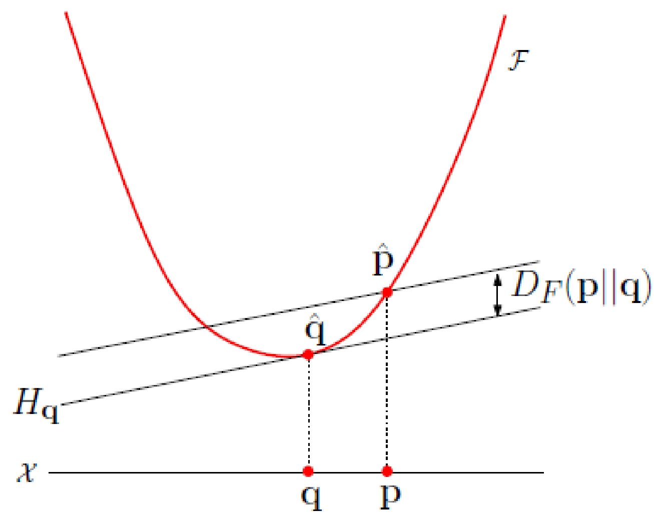


Figure 15: The univariate Bregman divergence interpretation as the vertical distance between the potential function φ and the hyperplane at H_q where F has been replaced by φ [95].

- The Euclidean distance is a Bregman divergence which is obtained from the generator $\varphi(p) = \sum_{i-1}^{d} p_i^2 = \langle p, p \rangle$, which is represented by the paraboloid function shown in figure 1
- The Kullback-Leibler divergence also is another Bregman divergence that is obtained from the generator $\varphi(p) = \sum_{i-1}^{d} p_i \log p_i$, which represents the negative entropy on a probability vector.

The Kullback-Leibler divergence is defined for simplex discrete distributions as:

$$KL(p||q) = \sum_{s} p(S=s) \log(\frac{p(S=s)}{q(S=s)}) \tag{5.4}$$

In general, $KL(p||q) \neq KL(q||p)$ is a non-symmetric divergence measurement. The KL divergence can be symmetric only if $p=q$. [92] proposed a way to symmetrize the KL SD_{φ} divergence between p and q, as below:

$$SD(p,q) = KL(p \parallel q) + KL(q \parallel p) \tag{5.5}$$

However, the SD_{φ} divergence can be undefined if q = 0 and p ≠ 0, which means that both values p and q should be continuous. In [93], a new algorithm was proposed to use the Jensen-Bregman divergence (JBD) to overcome these drawbacks:

$$JBD(p,q) = \frac{D_{\varphi}(p,\frac{q+p}{2}) + D_{\varphi}(q,\frac{q+p}{2})}{2}$$

$$= \frac{\varphi(p)+\varphi(q)}{2} - \varphi(\frac{p+q}{2}) \tag{5.6}$$

for d-dimensional multivariate data:

$$JBD(p,q) = \sum_{i=1}^{d} \frac{\varphi(p_i)+\varphi(q_i)}{2} - \varphi(\frac{p_i+q_i}{2}) \tag{5.7}$$

The geometric interpretation of the Jensen-Bregman divergence can be understood from the illustration in figure 16. In this figure, the vertical distance between the midpoint of the segment between $((\frac{p+q}{2}), \varphi(\frac{p+q}{2}))$ and the midpoint of $[(p,\varphi(p)),(q,\varphi(q))]$.

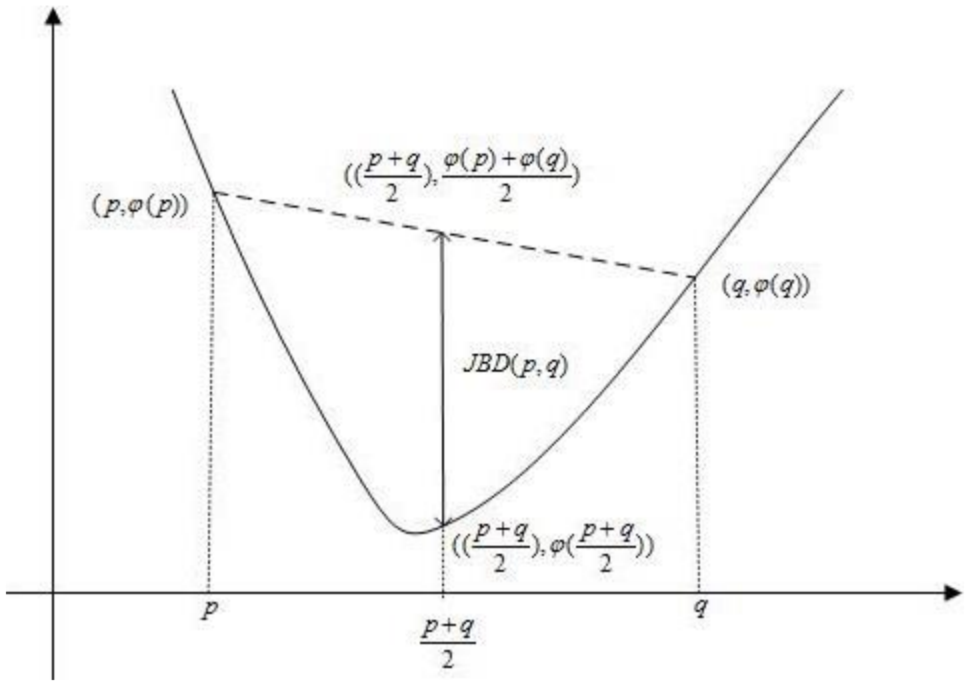


Figure. 16: Interpreting the Jensen-Bregman divergence.

All the quadratic distances $\varphi(p) = \langle Qp, p \rangle$ for a positive definite matrix, that are well known as squared Mahalanobis distance, can be found in the Jensen-Bregman divergence [93]:

$$
\begin{aligned}
& JBD(p,q) = \frac{\varphi(p)+\varphi(q)}{2} - \varphi(\frac{p+q}{2}) \\
& = \frac{2\langle Qp, p \rangle + 2\langle Qq, q \rangle - 2\langle Q(p+q), p+q \rangle}{4} \\
& = \frac{1}{4}(\langle Qp, p \rangle + \langle Qq, q \rangle - 2\langle Qp, q \rangle) \\
& = \frac{1}{4}\langle Q(p-q), p-q \rangle \\
& = \frac{1}{4}\|p-q\|_Q^2
\end{aligned}
$$

It is well known that the square root of the Jensen-Shannon divergence is a metric if the Shannon entropy generator has been used as $\varphi(p) = -p\log p$.

5.4 PNN-JBD IPS Method

The PNN is an implementation of a Kernel discriminate analysis algorithm that is derived from a Bayesian framework. PNN is organized into a four-layer feed forward network: input layer, pattern layer, summation layer and output layer. The first layer represents the input features. The input layer does not perform any computation. The second layer works as an activation function, the radial basis function (RBF). The prior probabilities in PNN are unknown. In order to estimate the prior probabilities, [100] proposes a probability kernel-based approach:

$$P(p,q) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(p-q)^2}{2\sigma^2}\right) \tag{5.8}$$

where σ is the kernel smoothing factor. The probability output will be equal to 1 if p = q, but the output will decay when the difference becomes larger between p and q. The architecture of PNN is shown in figure 17.

Input Layer: This layer consists of the multivariate input vector employed to estimate the location of the object. The number of neurons is equal to the number of the APs used in the area of interest.

Pattern (hidden) layer: In this layer the learning set will be organized to record the parameters of the input vector. The RPs in the training database were used to estimate the PDFs in order to estimate the likelihood of the input, so we can predict the class that the input vector belongs to, as shown in algorithm 1.

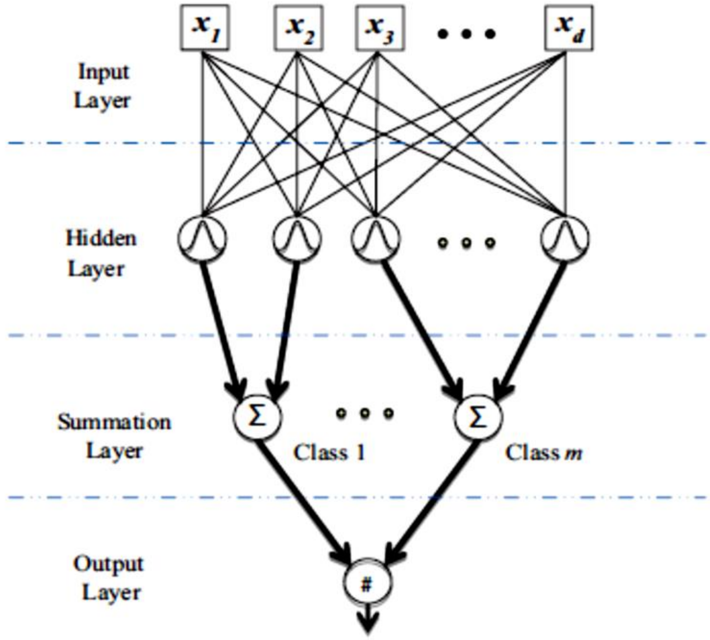


Figure 17. PNN Architecture [101].

Algorithm 1. The PNN positioning method

1- During offline phase, collect the RSS measurement from APs at specific locations to generate the fingerprinting map.
2- During online phase, collect the RSS from the APs at the unknown position, set the APs in the same way as the database of the offline phase with respect to the similar MAC address.
3- During online phase, perform the following steps for each fingerprint cell:
    - Estimate the smoothing factor σ that maximizes the output.
    - Estimate the similarity/dissimilarity using Eq. 4.8
4- Transfer the maximum outputs to the Output Layer.

Modified Training Algorithm: A new class of information-theoretic divergences that encapsulate both the Jensen-Shannon divergence and the squared Euclidean distance was used to estimate the probabilities from the training database using a JBD kernel-based approach to estimate the coordinate of the object. The RSS distribution will be defined as

$$kJBD(p,q_\ell) = \exp\left(-\frac{JBD(p(S_j\{x,y\}),q(S_j\{x_\ell,y_\ell\}))}{2\sigma^2}\right) \tag{5.9}$$

The PNN-JBD positioning method was proposed in order to improve the accuracy as shown in Algorithm 2:

Algorithm 2. The PNN-JSD positioning method

1- During offline phase, collect the RSS measurement from APs at specific locations to generate the fingerprinting map.
2- During online phase, collect the RSS from the APs at the unknown position, set the APs in the same way as the database of the offline phase with respect to the similar MAC address.
3- During online phase, perform the following steps for each fingerprint cell:
   - Estimate the smoothing factor σ that maximizes the output.
   - Estimate the similarity/dissimilarity using Eq. 4.9 and load the data into single column.
4- Transfer the maximum outputs to the Output Layer.

Output Layer: The number of neurons in this layer is equal to the number of existing classes. Each neuron must be connected to all the neurons in the hidden layer. As a result, this layer is also called the summation layer. The summation of the output layer is proportional to the density function as a Weighted Kernel Regression [84], that will be used to estimate the location:

$$(\bar{x}, \bar{y}) = \frac{\sum_{\ell} (x_{\ell}, y_{\ell}) kJBD(p, q_{\ell})}{\sum_{\ell} kJBD(p, q_{\ell})} \tag{5.10}$$

A regression model in equation 10 is used with K nearest neighbors as opposed to using the whole database in the hidden layer.

5.5 Implementation and Experimental Results

We implemented our algorithm inside the College of Engineering and Applied Sciences (CEAS) at Western Michigan University (WMU). In the first floor in CEAS, we used the layout shown in figure 18. This area has a long corridor and three rooms with a large study lounge in the area 23.5 m × 16.5 m. We employed an Android-based operating system 4.4.2 Samsung S5 smart phone as the test mobile to collect the RSS sampling. Furthermore, we implemented Java software by using an Eclipse framework version 4.2, which was installed on HP Pavilion for localization services. The APs were Cisco Linksys E2500 Advanced Simultaneous Dual-Band Wireless-N Routers in the area of interest.

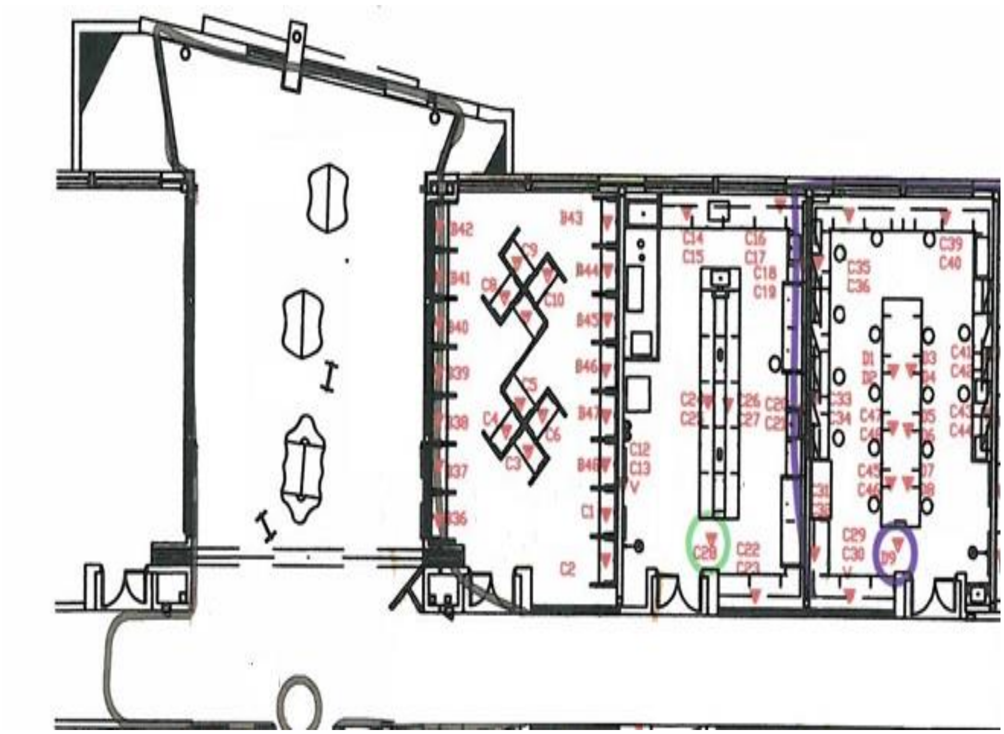


Figure 18: The layout used in the experimental work in the College of Engineering and Applied Sciences at WMU.

In the offline phase, the person that holds the phone to collect the data can increase the variation of the signal; furthermore, the number of passing individuals changes with time which also impacts the variation of the signal. In response, we created a realistic scenario that can provide a richer distribution for the WiFi signal. Therefore, the RSS recording was taken at four different orientations (45 °, 135 °, 225 °, and 315°) with ten scans and at a delay of 10 seconds taken at the same place to reduce the effects of the body and signal variations. The average value was taken for the four orientations and the ten recordings to generate the fingerprinting map.

The estimation accuracy can greatly be affected by the number of APs in the area of interest. In order to increase the accuracy of the positioning estimation, the impact of the number of the APs on the RPs were investigated in the estimation of the object position. To investigate the accuracy of our proposed algorithm, different algorithms have been implemented such as PNN and kNN as presented in [100] and [78] and compared with our proposed algorithm.

Different numbers of the nearest neighbors (NN) were used to estimate the coordination of the object. Figure 19 plots the localization distance with different numbers of APs when using 5 nearest neighbors only. The best performance was obtained when using 22 APs. The lowest distance error was 0.98m for PNN-JSD, 1.38m for kNN and 1.34m for PNN. Furthermore, the system showed higher accuracy when a kernel regression for 20 nearest neighbors was used; for example, the median accuracy for PNN-JBD was 0.92m, while 1.097m for PNN and 1.19m for kNN as shown in figure 20. Further decreases in distance error and better accuracy were observed when using 80 nearest neighbors for different AP selectins as shown in figure 21. The lowest median error distance was 0.865m for PNN-JSD, 0.99 m for PNN and 1.12m for kNN. Other positioning systems approaches such as compressive sensing [88], the kernel-based method [87], and kNN [78] reported results are less accuracy when compared to ours. Our proposed algorithm has shown lower distance error of maximum 90% of the other algorithms as shown in Table 3.

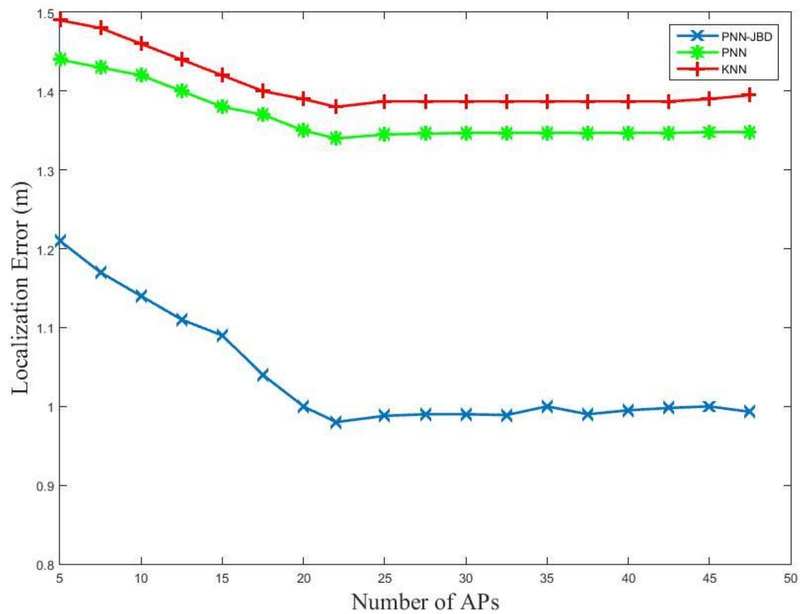


Figure 19: Localization distance error of PNN-JBD versus different algorithms with different number of APs with 5 NN

Table 3: Localization distance error of PNN-JBD versus different algorithms

Technique	Median [m]	Accuracy 90% [m]
kNN	1.8	3.7
Kernel-based	1.6	3.6
CS-based	1.5	2.7
PNN-JBD	0.899	1.85

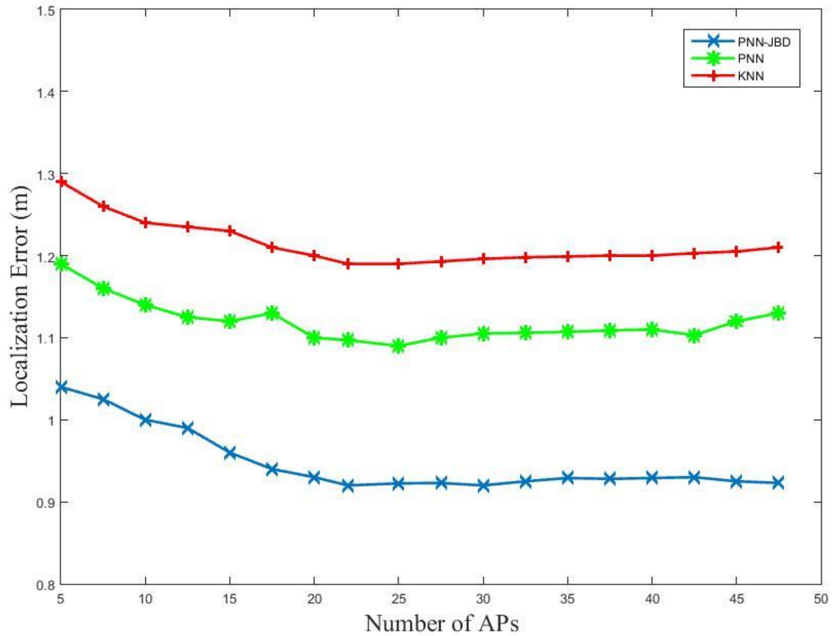


Figure 20: Localization distance error of PNN-JBD versus different algorithms with different number of APs with 20 NN

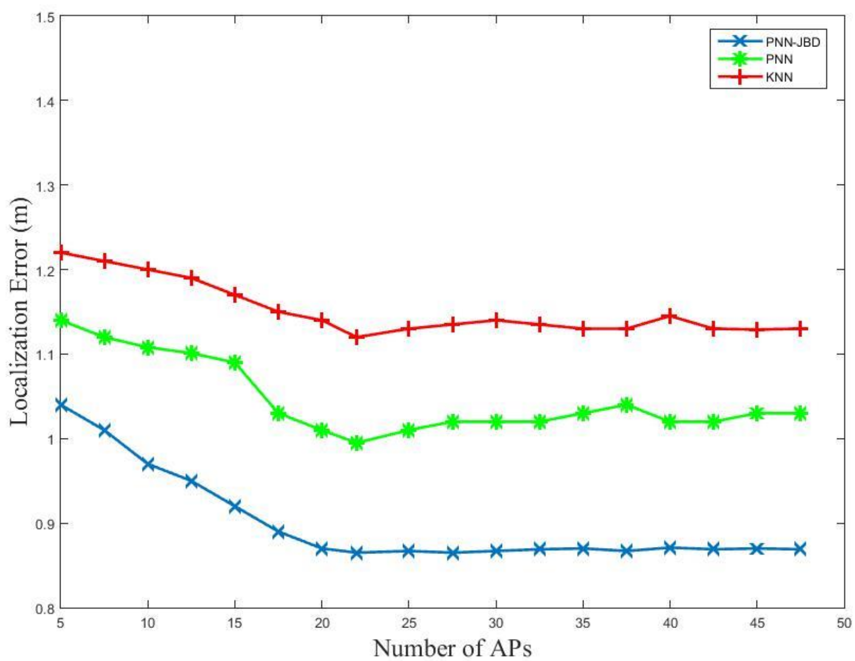


Figure 21: Localization distance error of PNN-JBD versus different algorithms with different number of APs with 80 NN

5.6 Conclusion

Indoor positioning systems have become a common tool in our daily life. A fingerprint-based scheme has become widely used in IPS. Furthermore, the fact that a large number of APs exist in our environment provides an easier way to investigate fingerprinting-based approaches. In this paper, we presented an indoor localization fingerprint-based scheme using a hybrid method composed of JBD and PNN. Experimental results were validated by comparing our proposed framework results with those of PNN and KNN methods. The PNN-JBD results are more accurate than the results of PNN and kNN. Also, we report that our best estimates occurred when incorporating 80 neighboring APS. We are in the process of investigating error modeling versus WiFi signal variation in space and time.

CHAPTER VI

A K-MEAN- JENSEN-SHANNON DIVERGENCE SYMMETRIZATION FOR A WLAN INDOOR POSITIONING SYSTEM

6.1 Introduction

Nowadays, smartphones have been rapidly proliferated, which has a lot of applications that are based on location-based services in pervasive computing and Internet of Things [117]. In general, GPS can't be used in indoor environments because the satellite signal can't penetrate the building, which makes it untraceable. Many technologies have been proposed and used instead of GPS that have different accuracy, such as ultra-wideband, Bluetooth, ultrasonic, and RSS. Most of the proposed methods consider the cost perspective, so the majority of the researchers depend on the existing infrastructure, or the deployment of dedicated infrastructure [unsupervised]. In the recent years, the RSS of wireless local access networks (WLAN) has become the dominant field of research in indoor positioning systems due to its availability and accuracy [119], and also the tremendous spread of chipsets from IEEE 802.11 that you can find both in smartphones and the APs in the area of interest [118]. There are two main categories in RSS-based indoor localization: fingerprinting and lateration-based techniques. In general, the lateration methods have less accuracy and more inaccurate estimation compared to fingerprinting methods because if one or more coordinates of the APs are not accurately estimated the performance can drop dramatically, and the non-line of sight (NLOS) problem that can affect the log distance propagation model (PM) [119]. While fingerprinting-based localization doesn't require LoS, the Wi-Fi fingerprinting is a process of collecting the WLAN signal with their indoor location

[120]– [123]. In general, each position in an indoor environment is characterized by a signal pattern, thus without knowing the AP location nor the angle measurement to estimate the object location. The fingerprinting-based technique is based on two phases: the offline phases (training) and the online phase as shown in figure 1. In the offline phase, a site survey will be performed to collect the RSS of the Wi-Fi from the different APs at known locations called reference points (RPs) in the area of interest. After that all the recorded signals will be stored at the database with their location for online query. While in the online phase, the user will send some samples of the RSS to the server to measure the similarity/dissimilarity using metric algorithms such as the Euclidean distance. The location of the object will be estimated based on the closest neighbors of the RPs set to match the target [119]. In this work, we propose:

- In order to obtain a richer characterization, the RSS was recorded into four different orientations (45 °, 135 °, 225 °, and 315°) to prevent a body blocking effect and reduce the signal variations.
- The Bregman k-means, that is the original k-means algorithm, is extended into a meta-algorithm will be used, since the Bregman divergence unifies the statistical entropic measures with the quadratic Euclidean distance to measure the similarity of the signals.
- The proposed algorithm results will be compared with the results of k-mean and affinity propagation with respect to accuracy and the average error distance.

6.2 Related Work

Nowadays, most of the research in Wi-Fi fingerprinting-based localization algorithms focuses on improving the collection of the fingerprinting signal database, which can lead to improve the accuracy and decrease the distance error estimation. Various algorithms have been proposed, some include ray-tracing [124], others use the signal propagation model, and some use crowdsourcing by using indoor floor maps and inertial sensors [125]. In general, the Wi-Fi signal suffers from time variation due to the nature of radio propagation which makes some difference between the offline and the online phase.

To eliminate the impact of the time variation, some researchers use clustering techniques by partitioning the fingerprinting database into multiple clusters and then choosing the one with the lowest average distance RSS to estimate the positioning of the target [117]. [126] proposed the cluster filtered KNN (CFK) method that partitions the fingerprint using hierarchical clustering; some improvements have shown in the results when clustering methods were used. Altintas and Serif [127] proposed an algorithm to replace the hierarchical clustering with k-mean to improve the accuracy of the positioning systems. Likewise, Sun et al. [128] proposed a KNN-FCM hybrid algorithm, a hybrid algorithm that uses the fuzzy *c*-means (FCM) clustering method incorporated with kNN of several clusters, and chose one cluster to estimate the object position. The results showed a little improvement with a distance error less than 2 meters. Tian [129] proposed an affinity propagation to cluster the fingerprinting database, after that coarse position algorithm that usually working with one cluster or more to estimate the location of the object.

In general affinity propagation clustering takes more time than the other techniques to cluster the database. [119] proposed a probabilistic neural network (PNN) scheme in which we incorporate the Jensen-Shannon divergence method.

6.3 Overall Structure of Proposed Positioning Algorithm

Designing a high accuracy IPS by depending on fingerprinting-based location is tricky because the RSS is heterogeneous. As a result, using the square Euclidean distance or Lp norm methods don't always give the highest accuracy results. For example, it has been proved in [119] that using the information theoretic relative entropy can obtain a better accuracy than the methods that depend on Euclidian distance such as PNN and kNN to measure the similarity between the offline phase and online phase. The Shannon differential entropy is defined as:

$$H(p) = p \log \frac{1}{p} \tag{6.1}$$

The kullback-Leibler divergence is equal to the cross entropy minus the Shannon differential entropy [95]:

$$KL(p||q) = \sum_{s} H^{x}(p(S=s)) \,\|\, (q(S=s)) - H(p(S=s)) \tag{6.2}$$

where H^{x} is the cross-entropy:

$$H^{x}(p(S=s) \,\|\, (q(S=s) = \sum_{s} (p(S=s) \log \frac{1}{(q(S=s)} \tag{6.3}$$

where S represents the vector sets of RSS. The Kullback- Leibler divergence of the two real-values p and q of the histogram distribution is defined as:

$$KL(p||q) = \sum_{s} p(S=s) \log(\frac{p(S=s)}{q(S=s)}) \tag{6.4}$$

Due to the hardware variance problem and the variation of the RSS, we use clustering methods to cluster the radio map of the offline phase. The k-mean was one of the first algorithms in clustering proposed by Lloyd in 1957 [131]. Briefly a k-mean iterative clustering algorithm was proposed to solve vector quantization problems. In general, k-mean works first by choosing a seed for each cluster considered as the cluster center, after that, the cluster center will associate the closest point to the center. This operation will be repeated and update the various cluster centers and will be reiterating and updated until the difference between any two successive calculations will be below the threshold. The cluster $C_{i's}$ center c_i is defined as follows [95]:

$$c_i = \arg\min \sum_{p_j \in c_i} \left\| p_j - c_i \right\| \tag{6.5}$$

$$= \arg\min\ AVG_{L_2^2}(C_i, c) \tag{6.6}$$

$$c_i = \frac{1}{|C_i|} \sum_{p_j \in c_i} p_j \tag{6.7}$$

where c_i represents the center of the cluster C_i, and $|C_i|$ represents the cardinality of C_i . In 2004 Banerjee et al. [130] proposed that the k-mean algorithm can be extended to a meta-algorithm by using a family of distortions called Bregman divergence. The Bregman divergence D_φ between two sets of convex data is defined as:

$$D_\varphi(p, q) = \varphi(p) - \varphi(q) - \langle \nabla \varphi(p), p - q \rangle \tag{6.8}$$

where $\nabla \varphi(p)$ denotes the gradient operator:

$$\nabla\varphi(p) = \left[\frac{\partial\varphi}{\partial p_1}\cdots\frac{\partial\varphi}{\partial p_d}\right]^T \tag{6.9}$$

and $\langle.,.\rangle$ denotes the dot product , φ is a strictly convex and differentiable function. Both The Euclidean distance and the Kullback-Leibler divergence can be derived from the Bregman divergence, if the convex function is considered as $\varphi(p) = \sum_{i-1}^{d} p_i^2 = \langle p, p\rangle$ that will lead to the Euclidean distance, while if the convex function is considered as $\varphi(p) = \sum_{i-1}^{d} p_i \log p_i$ that will lead to the Kullback-Leibler divergence as illustrated in figure 22.

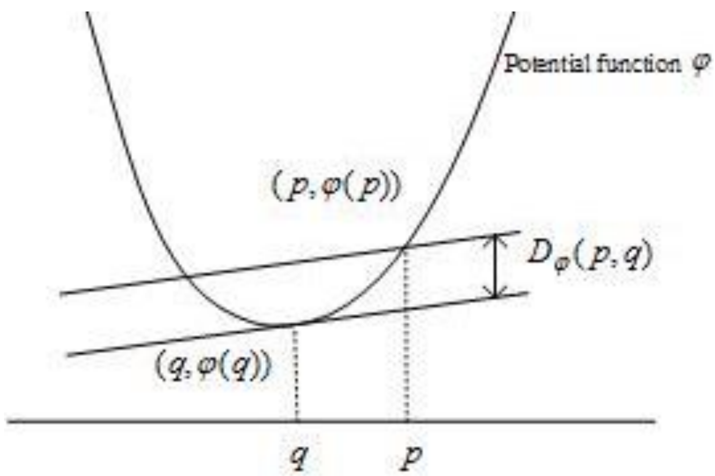


Figure 22: The univariate Bregman divergence interpretation as the vertical distance between the potential function φ and the hyperplane at q

The Bregman divergence can be symmetrized as follows:

$$SD_\varphi(p,q) = \frac{D_\varphi(p,q) + D_\varphi(q,p)}{2} \tag{6.10}$$

$$= \frac{1}{2}\left\langle p-q, \nabla\varphi(p) - \nabla\varphi(q)\right\rangle \tag{6.11}$$

The Kullback-Liebler divergence can be symmetrized in the same manner of the Bregman divergence as follows:

$$J(p,q) = KL(p \parallel q) + KL(q \parallel p) \tag{6.12}$$

The J-divergence has a major drawback: the output can be undefind if $p \neq 0$ and $q=0$. To overcome these drawbacks, [133] proposed a Jensen-Shannon divergence (JSD) by depending on the KL-divergence as:

$$JSD(p \parallel q) = \frac{1}{2}\left(KL\left(p, \frac{p+q}{2}\right) + KL\left(q, \frac{p+q}{2}\right)\right) \tag{6.13}$$

$$\begin{aligned} JSD(p \parallel q) &= \sum_s p(S=s) \log \frac{p(S=s)}{\frac{1}{2}q(S=s) + \frac{1}{2}p(S=s)} \\ &+ \sum_s q(S=s) \log \frac{q(S=s)}{\frac{1}{2}q(S=s) + \frac{1}{2}p(S=s)} \end{aligned} \tag{6.14}$$

Thus the original k-mean is modified into a meta-algorithm called Bregman k-mean. Barnerjee [132] proved that the mean is the minimizer of the clustering set of the expected Bregman divergence. The centroid of the point set can be defined as minimum average distance

$$c = \arg\min_p \frac{1}{n} \sum_i JSD(p, p_i) \tag{6.15}$$

$$c_R^F = \arg\min_{c \in RP} \frac{1}{n} \sum_{i=1}^{n} JSD(p_i \parallel c) \tag{6.16}$$

$$c_L^F = \arg\min_{c \in RP} \frac{1}{n} \sum_{i=1}^{n} JSD(c \parallel p_i) \tag{6.17}$$

$$c^F = \underset{c \in RP}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} \frac{JSD(c \parallel p_i) + JSD(p_i \parallel c)}{2} \tag{6.18}$$

where c_R^F and c_L^F represent the sided centroid (where L stands for left and R for right) and the centroid c^F represents the symmetrized Bregman centroid and n represents the number of RPs in the cluster [95]. The Bregman k-means positioning method was proposed in order to improve the accuracy.

Algorithm 1. The k- mean Bregman positioning method

1- Record The RSS measurement during the offline phase at known locations to create a database of fingerprinting maps.
2- Initialize the c centroid for each clustering group.
3- Calculate the distance between the RPs and the centroid using Eq. 19-20.
4- Assign each RP to the nearest centroid of the cluster.
5- Repeat step 2, 3, and 4 until there are no changes in each cluster centroid.
6- During online phase, perform the following steps for each fingerprint cell:
   - Estimate the distance between the RSS of unknown location and the centroid of each cluster using Euclidean distance.
   - Determine the kNN number of the cluster that RSS of online phase belong to.
7- Estimate the maximum outputs using eq. 19.

$$(\bar{x}, \bar{y}) = \frac{\sum_{\ell} (x_{\ell}, y_{\ell}) kJBD(p, q_{\ell})}{\sum_{\ell} kJBD(p, q_{\ell})} \tag{6.19}$$

6.4 Implementation and Experiment Results

The proposed algorithm is implemented inside the College of Engineering and Applied Sciences (CEAS) at Western Michigan University (WMU). The area has a three big rooms and large lounge for studying with a long corridor in the area 23.5 m × 16.5 m. An android operating system was used to collect the data and test the algorithm using a Java software by using an Eclipse framework version 4.2, that has been installed on HP Pavilion for localization estimation. The APs were Cisco Linksys E2500 Dual-Band in the area of interest. In the training (offline) phase the person that holding the phone during collecting the data may increase the variance of the signal; also the people that passing during the process of collecting data can play a role in signal variation. In response a realistic scenario was created to provide a better distribution of the Wi-Fi signal. The RSS signal was recorded at four different orientations (45 °, 135 °, 225 °, and 315°) with a delay 10 seconds for ten scan at each direction. After that an average value were taken for the different direction and the ten recording to generate the fingerprinting map.

In general, the accuracy can be affected by the number of APs that were used. The impact of the number of APs and the number of RPs have been investigated in the estimation process of the object. In order to investigate the accuracy of the proposed algorithms, different algorithms have been implemented and compared with our proposed algorithm such as k-mean and affinity propagation. Figure 23 illustrates the localization distance error with different numbers of APs when using 6 cluster with 5 nearest neighbors to estimate the location of the object. The highest accuracy obtained when a 22 APs was used, which was 1.063m for Bregman k- mean algorithms, 1.2175m for k-mean with probability function, and 1.2885m when k-mean with kNN algorithm. Furthermore, a better accuracy was obtained when a 15 nearest neighbor was used; for instance, the localization distance error was 0.98m for Bregman k-mean algorithm, 1.05m for k-mean with probability function, and 1.16m when k-mean with kNN algorithm as shown in figure 24.

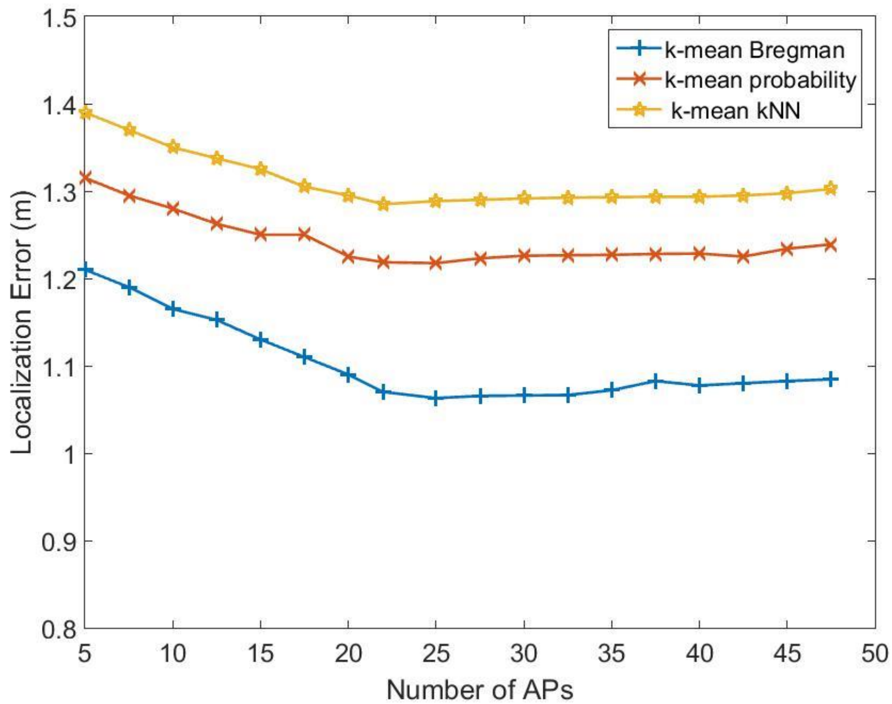


Figure 23: Localization distance error of k-mean Bregman versus different algorithms with different number of APs with 5 NN

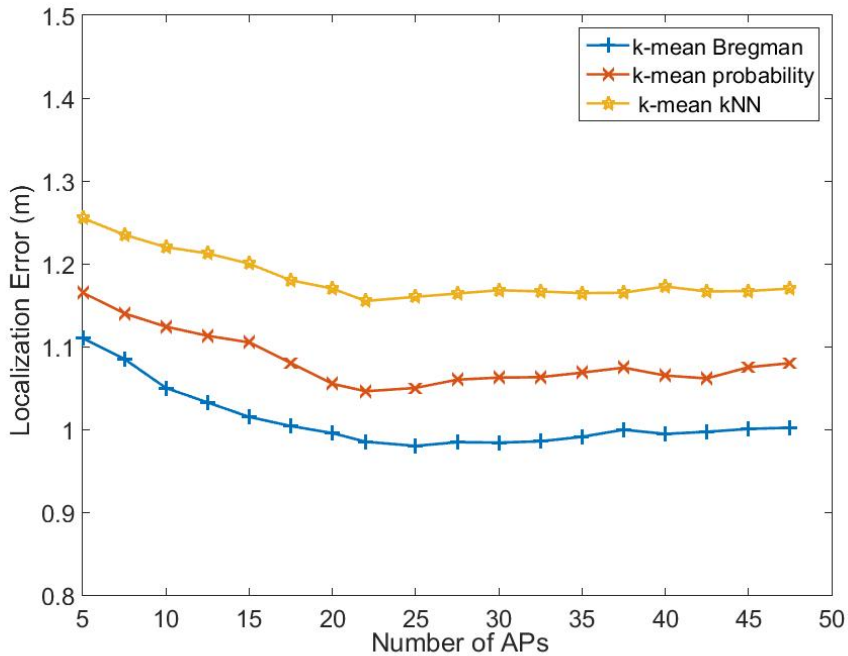


Figure 24: Localization distance error of k-mean Bregman versus different algorithms with different number of APs with 15 NN

There is some other positioning approach such affinity propagation [14] reported results are higher localization distance error. Furthermore, that our proposed algorithms showed higher accuracy of the other algorithms as shown in Table I.

Table 4: Localization distance error of k-mean Bregman versus different algorithms

Technique	**Median [m]**
k-mean+kNN	2.6
K-means + Prob. Dist	3.11

Affinity Prop. + Knn	2.66
Bregman k-mean	1.005

6.5 Conclusion

Indoor positioning systems are a very useful navigation tool in many applications in life. It can bring the power of the GPS indoors. In this paper, a WLAN positioning approach was proposed due to the tremendous number of APs in our environment provide an easier way to investigate the fingerprinting approach, called the Bregman k-means, that is the original k-means algorithm is extended into a meta-algorithm. The results that were obtained throughout our implementation showed that the Bregman k-mean outperforms the k-mean with kNN, k-mean with probability distribution, and the Affinity propagation algorithms. The best results were when 6 clustering was used with 15 NN. Nevertheless, now we are in the process to investigate the error modeling versus WiFi signal variation in space and time

CHAPTER VII

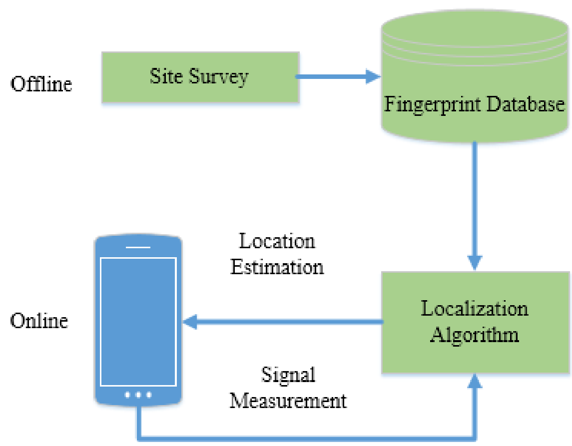
A JENSEN-BREGMAN DIVERGENCE FOR A WLAN INDOOR POSITIONING SYSTEM USING RECEIVED-SIGNAL-STRENGTH

7.1 Introduction

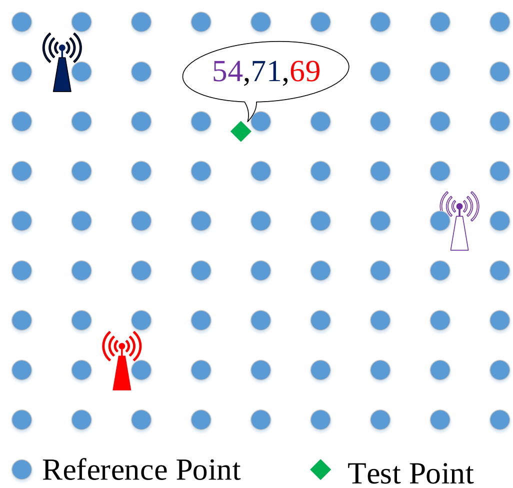
Nowadays, the automatic location of a user is a hot topic in research. [102] estimated the global indoor localization market around $935.05 million in 2014, and by 2019 it is expected to be around $4,424.1 million. The Compound Annual Growth Rate (CAGR) is expected to be 36.5% from 2014 to 2019. The estimation of mobile locations has an important role in many computing applications. In general, the Global Positioning System (GPS) is one of the most common location systems, but GPS cannot be used inside buildings since it can't perform a line-of-sight (LOS) with satellites and cannot determine the floor. Therefore, a large number of technologies were developed to create a high accuracy indoor positioning system (IPS): for example, Bluetooth, radio-frequency identification (RFID), wireless local area network (WLAN or Wi-Fi), magnetic field variations, ultrasound, ZigBee, LED light. Wi-Fi is the most common technique used in IPS. Because of the low cost, the existence of WLAN infrastructure and most of the smart phones can obtain the RSS from the access points (APs) of WLANs [103] [108].

The IPS algorithm that uses RSS-based indoor localization can be classified into two main types: the log-distance propagation model (PM) algorithms based on the signal, and fingerprinting indoor localization based on the data collection. The IPS based on signal propagation is divided into lateration and angulation. The main idea in lateration estimation is to calculate the distance between the smartphone and the access point (AP) by using geometry and signal measurement information, such as the time of arrival (TOA) of the signal from the APs, the time difference of arrival (TDOA) of the signals from the APs, and the angle of arrival (AOA) of the signals from the APs. In general, the propagation signal suffers from the non-line-of-sight (NLOS), multipath signal due to the walls, movement of people, and furniture. Also, the accuracy could be decreased if one or more coordinates of the APs haven't been accurately calculated. All these drawbacks made the estimation of the object using signal propagation a difficult task [104]. Therefore, to reduce the effect of these drawbacks, an implementation of the fingerprint-based signal has been proposed to estimate the location of the object [105]. Location fingerprinting was deployed because it doesn't require infrastructure, just the existing WLAN in the building and the smartphone, by depending on the characterization and spectrum of the RSS from the APs to the location to estimate the location coordinates.

The fingerprint-based technique has been divided into phases: the offline and online phases. In the offline phase the entire area of interest is divided into a rectangular set of grid points, and at each point, a site survey is taken by recording the RSS from the APs and stored in a database called a radio map. In the online phase, the smartphone will collect the RSS from APs and then send it to the server to compare the predefined fingerprint of the offline phase with the RSS in online phase to estimate the location on the grid map as shown in figure 25.



(a)



(b)

Figure 25: (a) The offline and online stages of location Wi-Fi based fingerprinting architecture and (b) radio map fingerprint of Wi-Fi IPS

kNN is one of the simplest ways to estimate the location by depending on the Euclidean distance to measure the similarity/dissimilarity between the offline and online phases. Even though this algorithm is easy to implement, it has low accuracy. Other methods like statistical learning and Bayesian modeling also have been used to estimate the location of the object. The location accuracy is one of the most fundamental metrics in IPS to measure the reliability of the system by reporting the error distance between the actual location and our estimated location [103].

Recently an important issue has been raised about the variation of the signal propagation: how can it change over time at the same place due to multiple factors such as physical obstructions, RF equipment, and the presence of human bodies? As a result, that can lead to attenuation and multipath issues, and this will make gradual changes in the signal which can reduce the accuracy of localization systems [106]. Values stored in the fingerprint maps represent the mean value of RSSI. Some approaches suppose that the RSSI distribution is a Gaussian [107] while other approaches assume a non-Gaussian, such as in [84]. However, using a Wi-Fi system to estimate the location of the object has many advantages, such as the availability and low cost to build a system compared to other technologies. But, on the other hand, there are some problems that we have to take into account, such as Wi-Fi hardware variance problems. Since the RSSI signal uses both off-line and on-line phases, this variance will affect the pattern of the signal, which will lead to the degradation of the accuracy of the location systems. Some experiments have been done to investigate this variance. It was reported in [106] that by using different smartphones to collect the RSSI data at the same time at the same location, some phones consistently reported a higher value of RSSI than the others. The orientation of the user can be a part of the variance of the RSSI signal because the human body can be a signification attenuator, as shown in figure 26, with the difference as much as 10 for same location different direction.

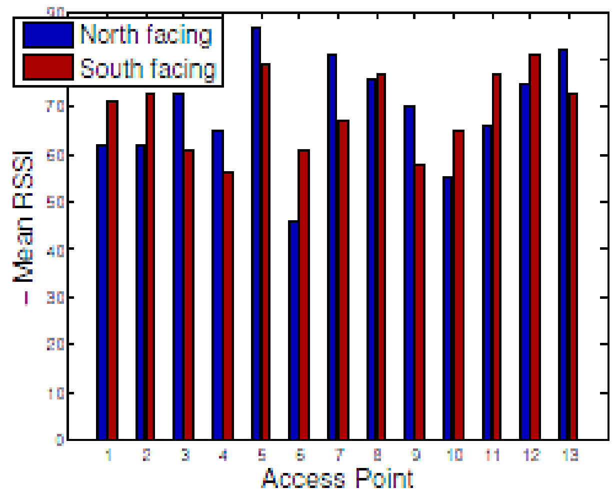


Figure 26: RSSI values recorded from different APS when facing North and South. The variation can be up to 10 dBm [106].

This variance hardware problem also has been noticed even in Cisco location systems, because some signal was omitted when a different device was used in the online phase than was used in the offline phase. In this work, we propose:

- A Jensen-Bregman Divergence (JBD) for a WLAN-based method and Kullback-Leibler Multivariate Gaussian KL_{MVG}. We perform the matching stage using probability kernels as a regression scheme.
- A procedure with high characterization distribution to be used. RSS value was taken in four different orientations (45 °, 135 °, 225 °, and 315°) to prevent body-blocking effects, with a scan for 100 seconds at each direction to reduce the effect of signal variation.
- JBD and KL_{MVG} results outperform the results of PNN and kNN with respect to accuracy and the average error distance, which indicates that the proposed

combining scheme is more effective in sensitive environments of WLAN-based positioning systems.

7.2 Related Work

The global navigation satellite systems (GNSS) like GLONASS, GALILEO or GPS work in the outdoor environment, but the accuracy may dramatically drop in indoor environments due to many parameters such as penetration loss, refraction, multipath propagation, and absorptions. Therefore, the necessity of developing a system that can work in an indoor environment with high accuracy has become imperative [108]. In the last decade, many techniques were proposed for IPS. In the model-based techniques, the location is estimated by depending on the geometrical model, such as in log-distance path loss (LDPL), where a semi-statistical function will be built by depending on the relationship between the RF propagation function and the RSS value. Several approaches have been proposed that are a trade-off between the accuracy and the cost, such as ToA, TDoA, AoA, and Multidimensional scaling (MDS). The MDS is a set of statistical techniques that is used to visualize the information in order to find the similarities/dissimilarities in the data. The matrix in MDS begins with item-item dissimilarities, the radio propagation attenuation between AP-AP to measure the distance [109]. The fingerprinting-based technique depends on matching algorithms, such as kNN, that have been used in RADAR [30], which is a pioneer of the fingerprint in IPS. After that, many developed kNN algorithms have been proposed to find the similarity/dissimilarity in metrics which usually is done by using the Manhattan distance or the Euclidean distance, such as in [110-112]. [113] proposed a new version of kNN that is more efficient than the probabilistic methods, Neural networks, and the traditional kNN, by depending on the decision tree of the training phases and taking the average of measures of RPs instead of having the whole dataset to estimate the location of the object. [114] performed a modified deterministic kNN technique with Mahalanobis, Manhattan, and Euclidian distances; their results showed that the Manhattan distance had higher accuracy than the others. Recently, the use of probabilistic distribution measurements in many IPS applications was increased. [98] Pioneered the use of the probabilistic distribution measurement in IPS. They propose a probabilistic framework by using a Bayesian network to estimate the location. In [36] a modified probability neural network (MPNN) was used to estimate the coordinates of the object. The results showed that the performance of MPNN is better than the triangulation methods. In [81] a kernel method was proposed to estimate the location of the object by using a histogram of the RSSI at the unknown location. In [82] the probability density function (PDF) was estimated by using the Kullback-Leibler divergence (KLD) framework as a composite hypothesis testing between the fingerprinting database and the test point. While in [83] they assumed that the RSSI distribution is multivariate Gaussian, and they used the KLD to estimate the impact of the RPs on the test point to estimate the probability of the closest one and then find the coordinates of the test point.

In [84] the RSS of the Bluetooth localization technique was used to establish the fingerprint and then the KLD was used in the probabilistic kernel regression to estimate the location of the object. The results showed around 1 m accuracy in an office environment. In general, the KLD kernel regression has better performance in multimodal distribution. In [89] the KLD was used to estimate the probabilistic kernel of both Gaussian and non-Gaussian distribution to compare between them and find their limitation.

5.3 Indoor Positioning System

We begin with a typical WLAN scenario, where a person carries a smartphone device that has a WLAN access, taking RSS measurements from the different APs at the College of Engineering and Applied Sciences (CEAS) at Western Michigan University (WMU). There is a common assumption that the RSSI coming from multiple APs is distributed as a multimodal signal as mentioned in [110]. However, the signal recorded quite different values for the same device at the same location, varying between two values different by as much as 10 dB. The values have been recorded for 35 minutes during rush hour for a single AP, for the same location, and then samples were taken from them as shown in figure 27.

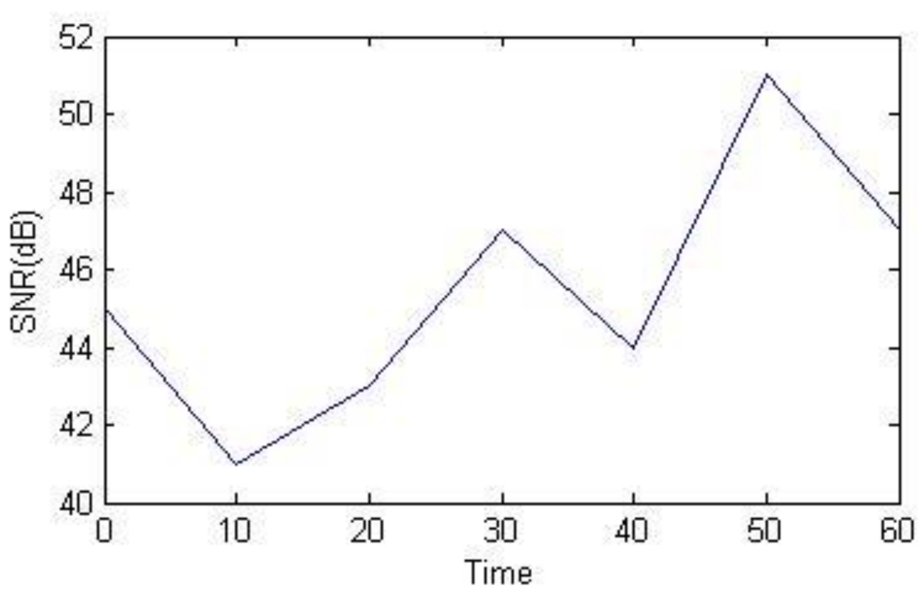


Figure 27: Signal-to-Noise Ratio (SNR) of the RSSI Variation Distribution over time

There are a lot of parameters that can affect the shape of the signal, like reflection and diffraction. Furthermore, the number of passing people have an impact on the shape of the signal. Nevertheless, we were looking for a scenario that could provide a better distribution of the Wi-Fi signal. During the offline phase, a realistic scenario was created that takes the variation of the signal into account, also, the effect of the body of person that holds the phone, and the passing of the people that can change the variation of the signal; however, in order to reduce the variation of the signal and the effect of the body, a recording of the RSS was taken in four directions (45 °, 135 °, 225 °, and 315°) . At each RP, a raw set of RSS was collected as a time sample from the APs in the area of interest that is denoted as $\{q_{i,j}^{(\circ)}(\tau), \tau = 1, ...,t, t = 100\}$, where t represents the number of time samples and $(\circ)$ is the orientation direction. After that, the average and covariance matrix of the RSS were taken for the four different directions and the ten scans to create the fingerprinting database, known as the radio map; the radio map is represented by $Q^{(\circ)}$ [86]:

$$Q^{(\circ)} = \begin{pmatrix} q_{1,1}^{(\circ)} & q_{1,2}^{(\circ)} & \cdots & q_{1,N}^{(\circ)} \\ q_{2,1}^{(\circ)} & q_{2,2}^{(\circ)} & \cdots & q_{2,N}^{(\circ)} \\ \vdots & \vdots & \ddots & \vdots \\ q_{L,1}^{(\circ)} & q_{L,2}^{(\circ)} & \cdots & q_{L,N}^{(\circ)} \end{pmatrix} \tag{7.1}$$

where $q_{i,j}^{(\circ)} = \frac{1}{q} \sum_{\tau=1}^{t} q_{i,j}^{(\circ)}(\tau)$, where t=10 and have been chosen randomly from the 100 time samples so we can obtain the average of the RSS sample over the time domain for different APs, $i = 1,2,.....L,\ j = 1,2,....N$, where N represents the number of RPs, and L is the number of the APs [49]. The variance vector of each RP can be defined as:

$$\Delta_j^{(\circ)} = \left[\Delta_{1,j}^{(\circ)}, \Delta_{2,j}^{(\circ)}, \Delta_{3,j}^{(\circ)}, ...\Delta_{L,j}^{(\circ)}\right] \tag{7.2}$$

where

$$\Delta_{i,j}^{(\circ)} = \frac{1}{t-1}\sum_{\tau=1}^{t}(q_{i,j}^{(\circ)}(\tau) - q_{i,j}^{(\circ)})^2 \tag{7.3}$$

where $\Delta_{i,j}^{(\circ)}$ is the variance for AP i at RP j with orientation $(\circ)$, so the database table of the radio map is ($x_j, y_j, q_j^{(\circ)}, \Delta_j^{(\circ)}$), and the $q_j^{(\circ)}$ will be defined as

$$q_j^{(\circ)} = \left[q_{1,j}^{(\circ)}, q_{2,j}^{(\circ)}, q_{3,j}^{(\circ)}, ..., q_{L,j}^{(\circ)}\right] \tag{7.4}$$

In the online phase, the RSS measurement will be denoted as:

$$p_r = \left[p_{1,r}, p_{2,r}, ..., p_{L,r}\right] \tag{7.5}$$

7.4 Kullback-Leibler Multivariate Gaussian Model

Recently another approach has been used in fingerprinting-based methods to estimate the position of the objects, the Multivariate Gaussian model (MvG), to exploit the interdependencies within the RPs, such as the model of the signal, the geometry that can be quantified to find the correlations among the RPs. Milioris [116] proposes a Kullback-Leibler multivariate Gaussian to measure the similarity between the RSS measurement of test points and the RPs that is defined as:

$$\mathrm{K}L_{MVG}(p \parallel q_j^{(\circ)}) = \frac{1}{2}\begin{pmatrix} (\mu_{q,j}^{S} - \mu_{p}^{S})^{T}(\Sigma_{j,q}^{s})^{-1} \\ (\mu_{q,j}^{S} - \mu_{p}^{S}) + tr(\Sigma_{R}^{s}(\Sigma_{j,q}^{s})^{-1} - I) \\ -\ln\left|\Sigma_{p}^{s}(\Sigma_{j,q}^{s})^{-1}\right| \end{pmatrix} \tag{7.6}$$

where S represents the matrix of RSS values from the different APs at specific locations, j represents the cell of the fingerprint location where:

$$S_j^{(\circ)} = \left\{\mu_j^{(\circ)}, \Sigma_j^{(\circ)}\right\} \tag{7.7}$$

$\mu_j^{(\circ)}$ the mean of *Jth* column of the RSS measurement, and $\Sigma_j^{(\circ)}$ represents the covariance matrix, where $|\Sigma|$ is the determinant of Σ. Now a probability kernel-based approach will be derived from the $\mathrm{K}L_{MVG}$. The Kernel regressions scheme allows us to estimate the PDF of the training datasets and the TP from the online phase that will be used to estimate the location of the object. The $\mathrm{K}L_{MVG}$ is used to measure the distance between the likelihood of the input sample and the RPs in order to find which class it belongs to. The RSS distribution can be defined as:

$$D(p, q_\ell) = \exp\left(-\frac{KL_{MVG}(p \parallel q_j^{(\circ)})}{2\sigma^2}\right) \tag{7.8}$$

where σ is the kernel smoothing factor. The probability will be equal to 1 if p = q, and the output will decrease when the difference between p and q becomes larger.

Algorithm 1. The Kullback-Leibler multivariate Gaussian positioning method

1. During the offline phase, the RSS measurement was taken at different places at know locations, ten scans with 10 seconds time delay to generate the radio map
2. During the online phase, the RSS measurement will be taken of the unknown location of the smartphone.
3. During online phase, the following steps will be performed:
   - A database for each RP will be set that have the RSS measurement with their location.
   - The RSS measurement from APs of the smartphone that have unknown location will be set in the same way as the database of the offline phase with respect to the similar MAC address.
   - Estimate the minimum Kullback-Leibler multivariate Gaussian using equation 7.8.
   - Repeat the step above for different APs until the minimum distance will be obtained.
4. Transfer the maximum outputs to the Output Layer.

7.5 Bregman Formulation

Analysis the data that is suffered from the interference and corrupted data is kind of impossible without interpreting the data that have been randomly obtained from unknown distribution with unknown parameters the most common assumption that have been used in many researches that the signal is Gaussian distribution. However, this assumption is inappropriate with RSS from the WLAN. The Gaussian distribution is considered as a member of the family of the exponentials distributions. Furthermore, the Bregman divergence and the exponential families have a strong relationship [133]. The log-likelihood of an exponential family will be considered as a sum of a Bregman divergence, However the Bregman divergence doesn't depend on the distribution parameter. The Bregman divergence can provide a likelihood distance of the exponential family, this property has been used to generalize the Principal Component Analysis (PCA) to the exponential family. However, the Bregman divergence is not a symmetric and doesn't satisfy the triangle inequality so it's not a metric. A Bregman divergence measures the distortion between classes that is defined by a Jensen convexity gap that is induced by a strictly convex function as:

$$D_{\varphi}(p,q) = \varphi(p) - \varphi(q) - \langle \nabla \varphi(p), p-q \rangle \tag{7.9}$$

where $\langle .,. \rangle$ denotes the inner product and

$$\langle p,q \rangle = \sum_{i=1}^{d} p^{(i)} q^{(i)} = p^T q \tag{7.10}$$

where $\nabla \varphi(p)$ denotes the gradient operator of φ at point q:

$$\nabla \varphi(p) = \left[\frac{\partial \varphi}{\partial p_1} \ldots\ldots \frac{\partial \varphi}{\partial p_d} \right]^T \tag{7.11}$$

The case of Bregman divergence is not a metric. However, as proved in section 3.2 that the Jensen-Bregamn divergence (JBD) is a symmetric and it can be a metric. JBD is induced by a strictly convex function generator that unifies the celebrated information-theoretic Jensen-Shannon divergence with the squared Euclidean and Mahalanobis distance:

$$JBD(p,q) = \sum_{i=1}^{d} \frac{\varphi(p_i)+\varphi(q_i)}{2} - \varphi(\frac{p_i+q_i}{2}) \tag{7.12}$$

The kernel function of JBD will be defined as:

$$kJBD(p,q_\ell) = \exp\left(-\frac{JBD(p(S_j\{x,y\}),q(S_j\{x_\ell,y_\ell\}))}{2\sigma^2}\right) \tag{7.13}$$

To improve the accuracy, we presented Algorithm 2:

Algorithm 2. The symmetric Bregman divergence positioning method

---

5. During the offline phase, the RSS measurement was taken at different places at known locations, ten scans with 10 seconds time delay to generate the radio map
6. During the online phase, the RSS measurement will be taken of the unknown location of the smart phone.
7. During online phase, the following steps will be performed:
    - A database for each RP will be set that have the RSS measurement with their location.
    - The RSS measurement from APs of the smart phone that has unknown location will be set in the same way as the database of the offline phase with respect to the similar MAC address.
    - Estimate the minimum symmetric Bregman divergence using 7.13 algorithm.
    - Repeat the step above for different APs until the minimum distance will be obtained.
8. Transfer the maximum outputs to the Output Layer.

7.6 Performance Analysis

The proposed algorithms evaluations will be demonstrated in the following subsections; the algorithms have been implemented in the first floor of the College of Engineering and Applied Sciences (CEAS) at Western Michigan University (WMU). This evaluation took place in an area 23.5 m × 16.5 m, with a large study lounge with three rooms and a long corridor as shown in figure 28. To collect the data sample a smartphone Samsung S5 with operating system 4.4.2 was employed. The proposed algorithms have been implemented on HP Pavilion by using Java software with an Eclipse framework. Cisco Linksys E2500 Advanced Simultaneous Dual-Band Wireless-N Routers were used in the area of interest. Most of the prior work ignored the variation of the RSS from the APs.

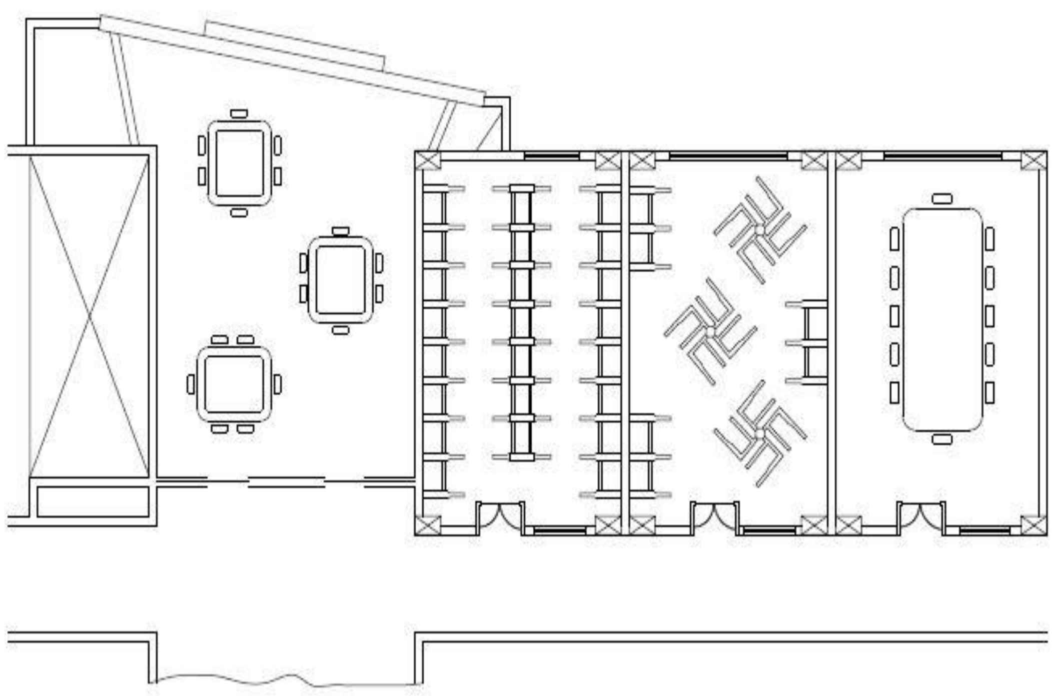


Figure 28: The layout used in the experimental work in the College of

Engineering and Applied

To evaluate the performance of the different fingerprinting techniques, the localization error was computed as the Euclidean distance between the actual reported coordinates of the test points and the coordinates of the mobile user during the online phase. The number of the RSS of the APs and the number of nearest neighbors have been noted that can affect the accuracy of the algorithms. It has been noticed that the number of the APs can play an important role in the accuracy of the distance error that can distinguish the near RPs from the other further away RPs.

In order to measure the impact of the APs on the accuracy, we used a specific number of nearest neighbors with a variety of APs. To investigate the accuracy of our proposed algorithm, different algorithms were used, such as PNN and KNN, and compared with our proposed algorithm. Different numbers of nearest neighbors (NN) were used to estimate the location of the object and evaluate the performance of our system framework. Figure 29 shows the impact of different APs when 5 NN were used. The lowest localization error was obtained when 22 APs were used which was 0.98m for kJBD, 1.12 m for kJSD, 1.16m for $\mathrm{K}L_{MVG}$, 1.34m for PNN and 1.38 m for kNN. More accuracy was obtained when more NN were used, as illustrated in figure 30, 22 NN was used. The lowest localization accuracy was obtained also when 22 APs were used as 0.92m for kJBD, 1.01m for *k*JSD, 1.02m for $\mathrm{K}L_{MVG}$, 1.097 m for PNN and 1.19 m for kNN. More improvement on system accuracy was noticed when 80 nearest neighbors were used: 0.865m for kJBD, 0.96 m for *k*JSD, 0.99 m for $\mathrm{K}L_{MVG}$, 0.995 m for PNN and 1.12 m for kNN as shown in figure 31. To validate our work, a comparison was made between the proposed algorithms with other algorithms from prior works such as kNN [10], compressive sensing [76] and the kernel-based method [77], illustrated in Table I.

Table 5: Localization distance error of different proposed algorithms

Technique	Median [m]	Accuracy 90% [m]
kNN	1.8	3.7
Kernel-based	1.6	3.6
CS-based	1.5	2.7
KL_{MVG}	1.02	2.13
kJSD	0.98	1.93

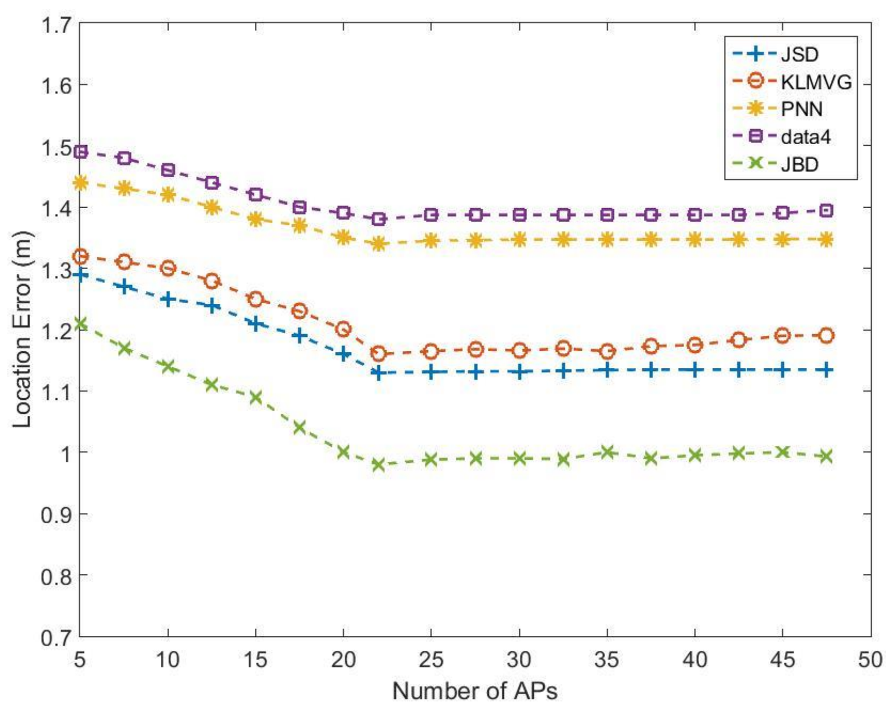


Figure 29: Localization distance error of different proposed algorithms with different number of APs with 5 NN

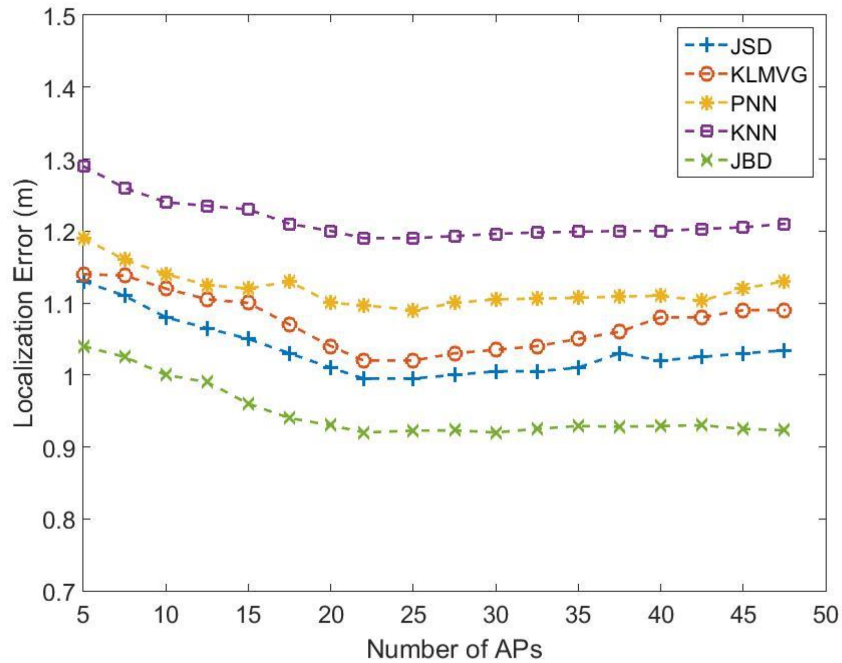


Figure 30: Localization distance error of different proposed algorithms with different number of APs with 20 NN

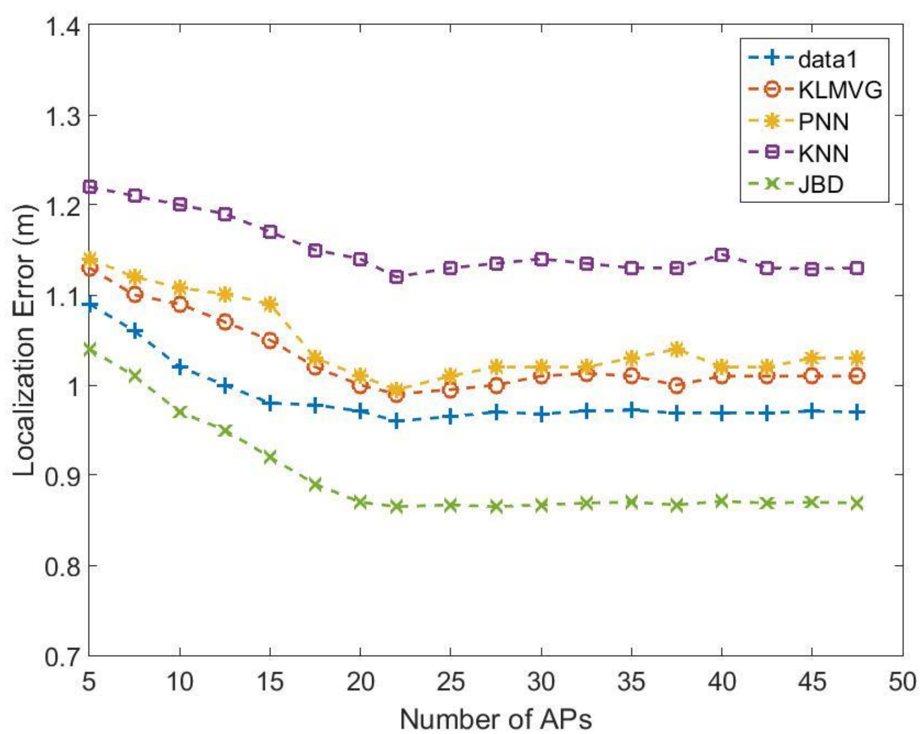


Figure 31: Localization distance error of different proposed algorithms with different number of APs with 80 NN

7.7 Conclusion

Indoor positioning systems bring the power of GPS and maps indoors. It can be a very useful navigation tool in many applications in life; for instance, emergency healthcare services, or for impaired vision people, or for use in unfamiliar buildings where people can get disoriented or lost easily, such as in mall, airport, subways. A fingerprint map was created for a segment of the college of engineering to utilize the relation of the RSS reading. Different algorithms were used and compared with different approaches such as kNN and PNN. The different performances were obtained for a number of the APs. The results were quite adequate for the indoor environment with an average error less than 1 meter. The kJBD had the highest accuracy when 80 NN with 22 APs among the other approach. Now we are in the process of investigating position prediction error distributions and need to quantify the localization variation of the WiFi signal distribution in space.

CHAPTER VIII

CONCLUSION AND FUTURE WORK

8.1 Summary and Conclusion

Fingerprint-based schemes have become widely proposed for indoor positioning systems. Furthermore, the fact that a large number of access points (AP) exist in indoor environments provides a convenient and economic context for investigating fingerprinting-based approaches.

The use of a Wi-Fi system to estimate the location of the object has many advantages such as its availability and lower cost when compared with other technology-based systems. On the other hand, it also presents certain problems that need to be taken into account such as Wi-Fi signal variation due to hardware differences and the very dynamic indoors environment. Since RSSI-based systems use both off-line and on-line phases, this variance would impact the accuracy of positioning system results. To address the issues of signal variation, the effect of the user's body and the interference caused by the presence of other people in the vicinity, signal acquisition was performed by recording the RSS in four directions (45 °, 135 °, 225 °, and 315°).

In this dissertation, several approaches were investigated by building on convex optimization models and the Bregman divergence was a natural tool to be selected. Since Bregman divergence family is well suited as distance measure tool, it was instinctive to combine these measures with probabilistic neural network (PNN) as a framework to measure the localization distance error. It was also as important to investigate Jensen-Bregman Divergence (JBD) since JBD is induced by a convex function generator and unifies the squared Euclidean and Mahalanobis distances with the information-theoretic Jensen-Shannon divergence. Finally, this investigation included a study of the number of the access points and location spacing versus localization accuracy.

The proposed algorithms were implemented using the College of Engineering and Applied Sciences (CEAS) at Western Michigan University (WMU) as the indoor environment testbed. A simulation of the multivariate Kullback-Leibler divergence (KL_{MVG}) under the Probability Neural Network (PNN) scheme and *k*-Nearest Neighbors (*k*-NN) was implemented to compare and allow for validation of the proposed framework using the Bregman divergence family. The JBD algorithm produced the smallest localization error and outperformed the other algorithms in terms of complexity and execution time. Therefore, the proposed framework, which is based on fingerprinting methods and convex optimization for minimization of the disparity measure, is a promising positioning system that is worthy of further exploration.

8.2 Future Work

Building on the proposed system, the following is a list of recommended future directions:

1. Pursue an investigation into the possibility of quantifying the variation in WiFi signal distribution versus space using clustering methods that allow the system to learn the number of clusters and the best number of nearest neighbors.

2. Automate data collection via a robot that can collect the data and develop a mechanism to keep the fingerprint map updated without any loss of service.
3. Develop better AP selection and feature extraction mechanisms to reduce computing complexity, storage needs, time, and effort.

BIBLIOGRAPHY

[1] Tian, Y. (2015). Practical Indoor Localization System Using GSM Fingerprints and Embedded Sensors, Universit_e Pierre et Marie Curie - Paris VI.

[2] NENA. 911 statistics. (2014). https://www.nena.org/?page=911Statistics.

[3] FCC. 911 wireless services. https://www.fcc.gov/guides/wireless-911-services.

[4] EENA. Position paper on the use of the GNSS to determine caller location in the context of 122 emergency call. (2014). http:// www.eena.org/ uploads/ gallery/files/ pdf/2014_05_EENA_Position_Paper_on_GNSS_full.pdf.

[5] Ingram, S., Harmer, D., & Quinlan, M. (2014). UltraWideBand indoor positioning systems and their use in emergencies. PLANS 2004. Position Location and Navigation Symposium (IEEE Cat. No.04CH37556). doi:10.1109/plans.2004.1309063.

[6] Strategy Analytics. http://www.strategyanalytics.com.

[7] Want, R., Hopper, A., Falcão, V., & Gibbons, J. (1992). The active badge location system. ACM Transactions on Information Systems, 10(1), 91-102. doi:10.1145/128756.128759.

[8] Hightower, J., & Borriello, G. (2001). Location systems for ubiquitous computing. Computer, 34(8), 57-66. doi:10.1109/2.940014.

[9] Küpper, A. (2005). Location-based services: Fundamentals and operation. Chichester, England: John Wiley.

[10] Dey, A. K. (2001). Understanding and Using Context. Personal and Ubiquitous Computing, 5(1), 4-7. doi:10.1007/s007790170019.

[11] Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggles, P. (1999). Towards a Better Understanding of Context and Context-Awareness. Handheld and Ubiquitous Computing Lecture Notes in Computer Science, 304-307. doi:10.1007/3-540-48157-5_29.

[12] Lukas, B. P. (2011). Robust Indoor Positioning through Adaptive Collaborative Labeling of Location Fingerprints. ETH, Zurich, 2011.

[13] Hahnel, D., Burgard, W., Fox, D., Fishkin, K., & Philipose, M. (2004). Mapping and localization with RFID technology. IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004. doi:10.1109/robot.2004.1307283.

[14] Harter, A., Hopper, A., Steggles, P., Ward, A., & Webster, P. (1999). The anatomy of a context-aware application. Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking - MobiCom '99. doi:10.1145/313451.313476.

[15] King, T., Haenselmann, T., & Effelsberg, W. (n.d.). Deployment, Calibration, and Measurement Factors for Position Errors in 802.11-Based Indoor Positioning Systems. Location- and Context-Awareness Lecture Notes in Computer Science, 17-34. doi:10.1007/978-3-540-75160-1_2.

[16] Gu, Y., Lo, A., & Niemegeers, I. (2009). A survey of indoor positioning systems for wireless personal networks. IEEE Communications Surveys & Tutorials, 11(1), 13-32. doi:10.1109/surv.2009.090103.

[17] Hightower, J., Brumitt, B., & Borriello, G. (2002). The location stack: A layered model for location in ubiquitous computing. Proceedings Fourth IEEE Workshop on Mobile Computing Systems and Applications. doi:10.1109/mcsa.2002.1017482.

[18] Krumm, J. (2008). A survey of computational location privacy. Personal and Ubiquitous Computing, 13(6), 391-399. doi:10.1007/s00779-008-0212-5.

[19] Frey, D., and Roman, G. (2007). Context-Aware Publish Subscribe in Mobile Ad Hoc Networks. Lecture Notes in Computer Science Coordination Models and Languages, 37-55. doi:10.1007/978-3-540-72794-1_3.

[20] Szymaszek, J., Uszok, A., & Zielinski, K. (1998). Building a scalable and efficient component-oriented system using CORBA - an Active Badge system case study. Distributed Systems Engineering, 5(4), 203-213. doi:10.1088/0967-1846/5/4/006.

[21] Steggles, P., and Gschwind, S. (2005). The ubisense smart space platform. Adjunct Proceedings of the 3rd International Conference on Pervasive Computing, pages 73-76.

[22] Chai, X., & Yang, Q. (2005). Reducing the Calibration Effort for Location Estimation Using Unlabeled Samples. Third IEEE International Conference on Pervasive Computing and Communications. doi:10.1109/percom.2005.34.

[23] Wan, K., & Alagar, V. (1997). Context-Aware Mobile Tour Guide. International Multi-Conference on Computing in the Global Information Technology - (ICCGI'06). doi:10.1109/iccgi.2006.1690323.

[24] A. Pourabdollah, X. Meng and M. Jackson, "Towards low-cost collaborative mobile positioning," Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2010, Kirkkonummi, 2010, pp. 1-5.

[25] Caffery, J. J. (1999). Wireless Location in CDMA Cellular Radio Systems. The Kluwer Academic Publishers, Norwell, MA, USA.

[26] Kim, W., Lee, J., & Jee, G. (2006). The Interior-Point Method for an Optimal Treatment of Bias in Trilateration Location. IEEE Transactions on Vehicular Technology, 55(4), 1291-1301. doi:10.1109/tvt.2006.877760.

[27] Alsindi, N., Alavi, B., & Pahlavan, K. (2007). Spatial Characteristics of UWB TOA-based Ranging in Indoor Multipath Environments. 2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications. doi:10.1109/pimrc.2007.4394445.

[28] Alsindi, N., Alavi, B., & Pahlavan, K. (2009). Measurement and Modeling of Ultrawideband TOA-Based Ranging in Indoor Multipath Environments. IEEE Transactions on Vehicular Technology, 58(3), 1046-1058. doi:10.1109/tvt.2008.926071.

[29] Dardari, D., Conti, A., Ferner, U., Giorgetti, A., & Win, M. Z. (2009). Ranging with Ultrawide Bandwidth Signals in Multipath Environments. Proceedings of the IEEE, 97(2), 404-426. doi:10.1109/jproc.2008.2008846.

[30] Irahhauten, Z., Bellusci, G., Janssen, G., Nikookar, H., & Tiberius, C. (2006). Investigation of UWB Ranging in Dense Indoor Multipath Environments. 2006 10th IEEE Singapore International Conference on Communication Systems. doi:10.1109/iccs.2006.301405.

[31] Lee, J., & Scholtz, R. (2002). Ranging in a dense multipath environment using an UWB radio link. IEEE Journal on Selected Areas in Communications, 20(9), 1677-1683. doi:10.1109/jsac.2002.805060.

[32] Zhang, D., Xia, F., Yang, Z., Yao, L., & Zhao, W. (2010). Localization Technologies for Indoor Human Tracking. 2010 5th International Conference on Future Information Technology. doi:10.1109/futuretech.2010.5482731.

[33] Li, W., Li, H., & Jiang, Y. (2014). A practical RF-based indoor localization system combined with the embedded sensors in smart phone. 2014 International Wireless Communications and Mobile Computing Conference (IWCMC). doi:10.1109/iwcmc.2014.6906522.

[34] Altini, M., Brunelli, D., Farella, E., & Benini, L. (2010). Bluetooth indoor localization with multiple neural networks. IEEE 5th International Symposium on Wireless Pervasive Computing 2010. doi:10.1109/iswpc.2010.5483748.

[35] Raghavan, A. N., Ananthapadmanaban, H., Sivamurugan, M. S., & Ravindran, B. (2010). Accurate mobile robot localization in indoor environments using bluetooth. 2010 IEEE International Conference on Robotics and Automation. doi:10.1109/robot.2010.5509232.

[36] Chawathe, S. S. (2009). Low-latency indoor localization using bluetooth beacons. 2009 12th International IEEE Conference on Intelligent Transportation Systems. doi:10.1109/itsc.2009.5309711.

[37] TI, CC2541 SensorTag development kit. http://www.ti.com/tool/cc2541dksensor.

[38] Mautz, R. (2012). Indoor positioning technologies. Zürich: Schweizerische Geodätische Kommission.

[39] Barnes, J., Rizos, C., Wang, J., Small, D., Voigt, G., & Gambale, N. (2003). High Precision Indoor and Outdoor Positioning using LocataNet. Journal of Global Positioning Systems, 2(2), 73-82. doi:10.5081/jgps.2.2.73.

[40] Jardak, N., & Samama, N. (2009). Indoor Positioning Based on GPS-Repeaters: Performance Enhancement using an Open Code Loop Architecture. IEEE Transactions on Aerospace and Electronic Systems, 45(1), 347-359. doi:10.1109/taes.2009.4805284.

[41] Moore, D., Leonard, J., Rus, D., & Teller, S. (2004). Robust distributed network localization with noisy range measurements. Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems - SenSys '04. doi:10.1145/1031495.1031502.

[42] Ye, R., & Liu, H. (2010). UWB TDOA localization system: Receiver configuration analysis. 2010 International Symposium on Signals, Systems and Electronics. doi:10.1109/issse.2010.5607083.

[43] Prorok, A., Tome, P., & Martinoli, A. (2011). Accommodation of NLOS for ultra-wideband TDOA localization in single- and multi-robot systems. 2011

International Conference on Indoor Positioning and Indoor Navigation. doi:10.1109/ipin.2011.6071927.

[44] Tarrio, P., Bernardos, A. M., Besada, J. A., & Casar, J. R. (2008). A new positioning technique for RSS-Based localization based on a weighted least squares estimator. 2008 IEEE International Symposium on Wireless Communication Systems. doi:10.1109/iswcs.2008.4726133.

[45] Yang, J., & Chen, Y. (2009). Indoor Localization Using Improved RSS-Based Lateration Methods. GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference. doi:10.1109/glocom.2009.5425237.

[46] Chintalapudi, K., Iyer, A. P., & Padmanabhan, V. N. (2010). Indoor localization without the pain. Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking - MobiCom '10. doi:10.1145/1859995.1860016.

[47] Subhan, F., Hasbullah, H., Rozyyev, A., & Bakhsh, S. T. (2011). Indoor positioning in Bluetooth networks using fingerprinting and lateration approach. 2011 International Conference on Information Science and Applications. doi:10.1109/icisa.2011.5772436.

[48] Ko, C. (2010). RFID 3D location sensing algorithms. Automation in Construction, 19(5), 588-595. doi:10.1016/j.autcon.2010.02.003.

[49] NOKIA. High accuracy indoor positioning. http://www.nomadicsolutions.biz/prod.php?lg=uk&id=51.

[50] Zhou, J., Zhang, H., & Mo, L. (2011). Two-dimension localization of passive RFID tags using AOA estimation. 2011 IEEE International Instrumentation and Measurement Technology Conference. doi:10.1109/imtc.2011.5944170.

[51] Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J. (2001). Global positioning system: Theory and practice," Springer, Wien, Austria.

[52] Diggelen, F. V. (2002). Indoor GPS theory & implementation. 2002 IEEE Position Location and Navigation Symposium (IEEE Cat. No.02CH37284). doi:10.1109/plans.2002.998914.

[53] Vittorini, L.D. and Robinson, B. (2003). Optimizing indoor GPS performance. GPS World, 14(11):40–48.

[54] Dovis, F., Lesca, R., Margaria, D., Boiero, G., & Ghinamo, G. (2008). An assisted high-sensitivity acquisition technique for GPS indoor positioning. 2008 IEEE/ION Position, Location and Navigation Symposium. doi:10.1109/plans.2008.4570123.

[55] Schon, S., & Bielenberg, O. (2008). On the capability of high sensitivity GPS for precise indoor positioning. 2008 5th Workshop on Positioning, Navigation and Communication. doi:10.1109/wpnc.2008.4510365.

[56] Zandbergen, P. A., & Barbeau, S. J. (2011). Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones. Journal of Navigation, 64(03), 381-399. doi:10.1017/s0373463311000051.

[57] Samama, N. (2008). Global positioning: Technologies and performance. Hoboken, NJ: Wiley-Interscience.

[58] Zimmermann, D., Baumann, J., Layh, M., Landstorfer, F., Hoppe, R., & Wolfle, G. (2004). Database correlation for positioning of mobile terminals in cellular networks using wave propagation models. IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004. doi:10.1109/vetecf.2004.1404980.

[59] Rehman, W. U., Lara, E. D., & Saroiu, S. (2008). CILoS: A CDMA indoor localization system. Proceedings of the 10th International Conference on Ubiquitous Computing - UbiComp '08. doi:10.1145/1409635.1409650.

[60] Buyruk, H., et al. (2013). RF fingerprinting based GSM indoor localization. 2013 21st Signal Processing and Communications Applications Conference (SIU). doi:10.1109/siu.2013.6531375.

[61] Borkowski, J., Niemela, J., & Lempiainen, J. (2004). Enhanced performance of cell ID RTT by implementing forced soft handover algorithm. IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004. doi:10.1109/vetecf.2004.1404724.

[62] Medbo, J., Siomina, I., Kangas, A., & Furuskog, J. (2009). Propagation channel impact on LTE positioning accuracy: A study based on real measurements of observed time difference of arrival. 2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications. doi:10.1109/pimrc.2009.5450144.

[63] Won, K., Jang, J., & Choi, H. (2012). An Enhanced Observed Time Difference of Arrival Based Positioning Method for 3GPP LTE System. IEICE

Transactions on Communications IEICE Trans. Commun., E95-B(3), 961-971.

[64] Bocquet, M., Loyez, C., & Benlarbi-Delaï, A. (2005). Millimeter wave up-converted UWB based positioning system. Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence Innovative Context-aware Services: Usages and Technologies - SOc-EUSAI '05. doi:10.1145/1107548.1107619.

[65] Gezici, S., Tian, Z., Giannakis, G., Kobayashi, H., Molisch, A., Poor, H., & Sahinoglu, Z. (2005). Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks. IEEE Signal Processing Magazine, 22(4), 70-84. doi:10.1109/msp.2005.1458289.

[66] Waadt, A., Wang, S., Kocks, C., Burnic, A., Xu, D., Bruck, G., & Jung, P. (2010). Positioning in multiband OFDM UWB utilizing received signal strength. 2010 7th Workshop on Positioning, Navigation and Communication. doi:10.1109/wpnc.2010.5653193.

[67] Montaser, A., & Moselhi, O. (2012). RFID indoor location identification for construction projects. Gerontechnology, 11(2). doi:10.4017/gt.2012.11.02.156.00.

[68] Azzouzi, S., Cremer, M., Dettmar, U., Knie, T., & Kronberger, R. (2011). Improved AoA based localization of UHF RFID tags using spatial diversity. 2011 IEEE International Conference on RFID-Technologies and Applications. doi:10.1109/rfid-ta.2011.6068634.

[69] Ozdenizci, B., Ok, K., Coskun, V., & Aydin, M. N. (2011). Development of an Indoor Navigation System Using NFC Technology. 2011 Fourth International Conference on Information and Computing. doi:10.1109/icic.2011.53.

[70] Ilkovičová, L., & Kopáčik, A. (2015). Development of an indoor navigation system using NFC technology. Advances and Trends in Engineering Sciences and Technologies Proceedings of the International Conference on Engineering Sciences and Technologies, 27-29 May 2015, Tatranská Štrba, High Tatras Mountains - Slovak Republic, 275-280. doi:10.1201/b19249-49.

[71] Ferro, E., & Potorti, F. (2005). Bluetooth and wi-fi wireless protocols: A survey and a comparison. IEEE Wireless Communications, 12(1), 12-26. doi:10.1109/mwc.2005.1404569.

[72] Chen, Y., Francisco, J., Trappe, W., & Martin, R. (2006). A Practical Approach to Landmark Deployment for Indoor Localization. 2006 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks. doi:10.1109/sahcn.2006.288441.

[73] Hu, T. (2013). Towards practical deployment of Wi-Fi fingerprint-based indoor localization. doi:10.14711/thesis-b1514761.

[74] Wang, J., Gao, Q., Yu, Y., Wang, H., & Jin, M. (2012). Toward Robust Indoor Localization Based on Bayesian Filter Using Chirp-Spread-Spectrum Ranging. IEEE Transactions on Industrial Electronics, 59(3), 1622-1629. doi:10.1109/tie.2011.2165462.

[75] Cho, H., & Kim, S. W. (2010). Mobile Robot Localization Using Biased Chirp-Spread-Spectrum Ranging. IEEE Transactions on Industrial Electronics, 57(8), 2826-2835. doi:10.1109/tie.2009.2036633.

[76] Dhital, A., Closas, P., & Fernández-Prades, C. (2012). Bayesian filtering for indoor localization and tracking in wireless sensor networks. EURASIP J Wirel Commun Netw EURASIP Journal on Wireless Communications and Networking, 2012(1), 21. doi:10.1186/1687-1499-2012-21.

[77] Talvitie, J., Renfors, M., & Lohan, E. S. (2015). Distance-Based Interpolation and Extrapolation Methods for RSS-Based Localization with Indoor Wireless Signals. IEEE Transactions on Vehicular Technology, 64(4), 1340-1353. doi:10.1109/tvt.2015.2397598.

[78] Bahl, P., & Padmanabhan, V. (2002). RADAR: An in-building RF-based user location and tracking system. Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064). doi:10.1109/infcom.2000.832252.

[79] Wan, Q., Yuan, J., Peng, Y., & Long, K. (2003). A quadrant-based location tracking method. Proceedings of the 2003 IEEE Radar Conference (Cat. No. 03CH37474).

[80] Castro, P., Chiu, P., Kremenek, T., & Muntz, R. (2001). A Probabilistic Room Location Service for Wireless Networked Environments. Ubicomp 2001: Ubiquitous Computing Lecture Notes in Computer Science, 18-34. doi:10.1007/3-540-45427-6_3.

[81] Roos, T. et al. (2002). A Probabilistic Approach to WLAN User Location Estimation. International Journal of Wireless Information Networks, vol.7, n.3, 2002.

[82] Tsui, A. W., Chuang, Y., & Chu, H. (2009). Unsupervised Learning for Solving RSS Hardware Variance Problem in WiFi Localization. Mobile Networks and Applications, 14(5), 677-691. doi:10.1007/s11036-008-0139-0

[83] Milioris, D., Kriara, L., Papakonstantinou, A., Tzagkarakis, G., Tsakalides, P., & Papadopouli, M. (2010). Empirical evaluation of signal-strength fingerprint positioning in wireless LANs. Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems - MSWIM '10. doi:10.1145/1868521.1868525.

[84] Mirowski, P., Steck, H., Whiting, P., Palaniappan, R., Macdonald, M., & Ho, T. K. (2011). KL-divergence kernel regression for non-Gaussian fingerprint based localization. 2011 International Conference on Indoor Positioning and Indoor Navigation. doi:10.1109/ipin.2011.6071928.

[85] Wang, B., Zhou, S., Liu, W., & Mo, Y. (2015). Indoor Localization Based on Curve Fitting and Location Search Using Received Signal Strength. *IEEE Transactions on Industrial Electronics, 62*(1), 572-582. doi:10.1109/tie.2014.2327595.

[86] Chen, Y., Francisco, J., Trappe, W., & Martin, R. (2006). A Practical Approach to Landmark Deployment for Indoor Localization. *2006 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*. doi:10.1109/sahcn.2006.288441.

[87] Kushki, A., Plataniotis, K. N., & Venetsanopoulos, A. N. (2007). Kernel-Based Positioning in Wireless Local Area Networks. *IEEE Transactions on Mobile Computing, 6*(6), 689-705. doi:10.1109/tmc.2007.1017.

[88] Feng, C., Au, W. S., Valaee, S., & Tan, Z. (2012). Received-Signal-Strength-Based Indoor Positioning Using Compressive Sensing. *IEEE Transactions on Mobile Computing IEEE Trans. on Mobile Comput., 11*(12), 1983-1993. doi:10.1109/tmc.2011.216.

[89] Mirowski, P., Whiting, P., Steck, H., Palaniappan, R., Macdonald, M., Hartmann, D., & Ho, T. K. (2012). Probability kernel regression for WiFi localisation. *Journal of Location Based Services, 6*(2), 81-100. doi:10.1080/17489725.2012.694723.

[90] Misikangas,P. . Lekman, L. (2003). Applications of signal quality observations. Patent Cooperation Treaty (PCT) WO 2004/008796 A1.

[91] Ullah, A. (1996). Entropy, divergence and distance measures with econometric applications. *Journal of Statistical Planning and Inference, 49*(1), 137-162. doi:10.1016/0378-3758(95)00034-8.

[92] Boissonnat, J., Nielsen, F., & Nock, R. (2010). Bregman Voronoi Diagrams. *Discrete & Computational Geometry, 44*(2), 281-307. doi:10.1007/s00454-010-9256-1.

[93] Nielsen, F., & Nock, R. (2011). Skew Jensen-Bregman Voronoi Diagrams. *Lecture Notes in Computer Science Transactions on Computational Science XIV*, 102-128. doi:10.1007/978-3-642-25249-5_4.

[94] Wu, L., Huang, X., Zhang, C., Shepherd, J., & Wang, Y. (2014). An efficient framework of Bregman divergence optimization for co-ranking images and tags in a heterogeneous network. *Multimed Tools Appl Multimedia Tools and Applications, 74*(15), 5635-5660. doi:10.1007/s11042-014-1873-x.

[95] Nielsen, F., & Nock, R. (2008). Bregman sided and symmetrized centroids. *2008 19th International Conference on Pattern Recognition*. doi:10.1109/icpr.2008.4761794.

[96] Lotfi, A., & Benyettou, A. (2014). A reduced probabilistic neural network for the classification of large databases. *Turk J Elec Eng & Comp Sci TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES, 22*, 979-989. doi:10.3906/elk-1209-35.

[97] Majtey, A. P., Lamberti, P. W., & Prato, D. P. (2005). Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states. *Physical Review A, 72*(5). doi:10.1103/physreva.72.052310.

[98] Wu, C., Yang, Z., Liu, Y., & Xi, W. (2012). WILL: Wireless indoor localization without site survey. *2012 Proceedings IEEE INFOCOM*. doi:10.1109/infcom.2012.6195809.

[99] Cho, H., & Kim, S. W. (2010). Mobile Robot Localization Using Biased Chirp-Spread-Spectrum Ranging. *IEEE Transactions on Industrial Electronics*, *57*(8), 2826-2835. doi:10.1109/tie.2009.2036633.

[100] Chen, C., Yin, L., Chen, Y., & Hwang, R. (2012). A modified probability neural network indoor positioning technique. *2012 International Conference*

on Information Security and Intelligent Control. doi:10.1109/isic.2012.6449770.

[101] Mishra1, M. et al. (2013). A Probabilistic Neural Network Approach for Classification of Vehicle. International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 2, Issue 7, July 2013.

[102] Markets, & Markets. (2014). Indoor localization Market by Positioning Systems, Map and Navigation, Location based analysis, Monitoring and emergency Services-Worldwide Market Forecasts and Analysis (2014-2019. Technical Report, 2014.

[103] Torres-Sospedra, J., Montoliu, R., Trilles, S., Belmonte, Ó, & Huerta, J. (2015). Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. *Expert Systems with Applications, 42*(23), 9263-9278. doi:10.1016/j.eswa.2015.08.013.

[104] Shchekotov, M. (2015). Indoor localization methods based on Wi-Fi lateration and signal strength data collection. *2015 17th Conference of Open Innovations Association (FRUCT)*. doi:10.1109/fruct.2015.7117991.

[105] Swangmuang, N., & Krishnamurthy, P. (2008). An effective location fingerprint model for wireless indoor localization. *Pervasive and Mobile Computing, 4*(6), 836-850. doi:10.1016/j.pmcj.2008.04.005.

[106] Sharma, P., Chakraborty, D., Banerjee, N., Banerjee, D., Agarwal, S. K., & Mittal, S. (2014). KARMA: Improving WiFi-based indoor localization with dynamic causality calibration. *2014 Eleventh Annual IEEE International*

Conference on Sensing, Communication, and Networking (SECON). doi:10.1109/sahcn.2014.6990331.

[107] Ferris, B., Haehnel, D., & Fox, D. (2006). Gaussian Processes for Signal Strength-Based Location Estimation. *Robotics: Science and Systems II*. doi:10.15607/rss.2006.ii.039.

[108] Jiang, P., Zhang, Y., Fu, W., Liu, H., & Su, X. (2015). Indoor Mobile Localization Based on Wi-Fi Fingerprint's Important Access Point. *International Journal of Distributed Sensor Networks, 2015*, 1-8. doi:10.1155/2015/429104.

[109] Yang, Z., Wu, C., & Liu, Y. (2012). Locating in fingerprint space: Wireless indoor localization with little human intervention.*Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, 269-280.

[110] Youssef, Moustafa, Agrawala, Ashok, Kotz, David, Noble, Brian, & Shin, Kang, G. (2005). The Horus WLAN location determination system. *Mobile Systems, Applications, and Services Proceedings of the 3rd International Conference*, 205-218.

[111] Ni, Lm, Liu, Yh, Lau, Yc, & Patil, Ap. (2004). LANDMARC: Indoor location sensing using active RFID. *10*(6), 701-710.

[112] Sen, Souvik, Radunovic, Božidar, Choudhury, Romit, Roy, Minka, Tom, Davies, Nigel, Seshan, Srinivasan, & Zhong, Lin. (2012). You are facing the Mona Lisa: Spot localization using PHY layer information. *Mobile*

Systems, Applications, and Services Proceedings of the 10th International Conference, 183-196.

[113] Yim, J. (2008). Introducing a decision tree-based indoor positioning technique. *Expert Systems With Applications, 34*(2), 1296-1302.

[114] Farshad, A., Li, J., Marina, M. K., & Garcia, F. J. (2013). A microscopic look at WiFi fingerprinting for indoor mobile phone localization in diverse environments. *International Conference on Indoor Positioning and Indoor Navigation*. doi:10.1109/ipin.2013.6817920.

[115] Lin, J. (1991). Divergence measures based on the Shannon entropy. *Information Theory, IEEE Transactions on, 37*(1), 145-151.

[116] Milioris, Tzagkarakis, Papakonstantinou, Papadopouli, & Tsakalides. (2014). Low-dimensional signal-strength fingerprint-based positioning in wireless LANs. *Ad Hoc Networks, 12*, 100-114.

[117] Hu, X., Shang, J., Gu, F., & Han, Q. (2015). Improving Wi-Fi Indoor Positioning via AP Sets Similarity and Semi-Supervised Affinity Propagation Clustering. *International Journal of Distributed Sensor Networks, 2015*, 11.

[117] Pajovic, M., Orlik, P., Koike-Akino, T., Kim, K. J., Aikawa, H., & Hori, T. (2014). An Unsupervised Indoor Localization Method Based on Received Signal Strength (RSS) Measurements. *2015 IEEE Global Communications Conference (GLOBECOM)*. doi:10.1109/glocom.2014.7417708.

[119] Abdullah, O., Abdel-Qader, I., & Bazuin, B. (2016). A probability neural network-Jensen-Shannon divergence for a fingerprint based localization. *2016 Annual Conference on Information Science and Systems (CISS)*. doi:10.1109/ciss.2016.7460516.

[120] He, S., & Chan, S. G. (2016). Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons. *IEEE Communications Surveys & Tutorials, 18*(1), 466-490. doi:10.1109/comst.2015.2464084.

[121] Liu, H., Gan, Y., Yang, J., Sidhom, S., Wang, Y., Chen, Y., & Ye, F. (2012). Push the limit of WiFi based localization for smartphones. *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking - Mobicom '12*. doi:10.1145/2348543.2348581.

[122] Sun, W., Liu, J., Wu, C., Yang, Z., Zhang, X., & Liu, Y. (2013). MoLoc: On Distinguishing Fingerprint Twins. *2013 IEEE 33rd International Conference on Distributed Computing Systems*. doi:10.1109/icdcs.2013.41.

[123] Xiao, Z., Wen, H., Markham, A., Trigoni, N., Blunsom, P., & Frolik, J. (2015). Non-Line-of-Sight Identification and Mitigation Using Received Signal Strength. *IEEE Transactions on Wireless Communications IEEE Trans. Wireless Commun., 14*(3), 1689-1702. doi:10.1109/twc.2014.2372341.

[124] Noh, Y., Yamaguchi, H., Lee, U., Vij, P., Joy, J., & Gerla, M. (2013). CLIPS: Infrastructure-free collaborative indoor positioning scheme for time-critical team operations. *2013 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. doi:10.1109/percom.2013.6526729.

[125] Radu, V., & Marina, M. K. (2013). HiMLoc: Indoor smartphone localization via activity aware Pedestrian Dead Reckoning with selective crowdsourced WiFi fingerprinting. *International Conference on Indoor Positioning and Indoor Navigation*. doi:10.1109/ipin.2013.6817916.

[126] Ma, J., Li, X., Tao, X., & Lu, J. (2008). Cluster filtered KNN: A WLAN-based indoor positioning scheme. *2008 International Symposium on a World of Wireless, Mobile and Multimedia Networks*. doi:10.1109/wowmom.2008.4594840.

[127] Altintas, B. and Serif, T. (2011). Improving RSS-based indoor positioning algorithm via K-Means clustering. in Proceedings of the 11th European Wireless Conference 2011—Sustainable Wireless Technologies (EuropeanWireless), pp. 1–5, VDE, 2011.

[128] Yongliang Sun, Yubin Xu, Lin Ma, & Zhian Deng. (2009). KNN-FCM hybrid algorithm for indoor location in WLAN.*Power Electronics and Intelligent Transportation System (PEITS), 2009 2nd International Conference on*, *2*, 251-254.

[129] Nielsen, F., & Nock, R. (2009). Sided and Symmetrized Bregman Centroids. *Information Theory, IEEE Transactions on*, *55*(6), 2882-2904.

[130] Banerjee, A., Merugu, S., Dhillon, I., & Ghosh, J. (2005). Clustering with Bregman divergences. *Journal Of Machine Learning Research*, *6*, 1705-1749.

[131] Lloyd, S. (1982). Least squares quantization in PCM.*Information Theory, IEEE Transactions on*, *28*(2), 129-137.

[132] Pahlavan, Xinrong Li, & Makela. (2002). Indoor geolocation science and technology. *Communications Magazine, IEEE,40*(2), 112-118.

[133] Chen, P., Chen, Y., & Rao, M. (2008). Metrics defined by Bregman Divergences. Communications in Mathematical Sciences, 6(4), 915-926. doi:10.4310/cms.2008.v6.n4.a6