©2010

Akshay Shetty

ALL RIGHTS RESERVED

Weighted K-Nearest Neighbor Algorithm as an Object Localization technique

using Passive RFID Tags

by

AKSHAY SHETTY

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Electrical and Computer Engineering

written under the direction of

Professor. Ivan Marsic

and approved by

____

____

____

____

New Brunswick, New Jersey

October, 2010

ABSTRACT OF THE THESIS

Weighted K-Nearest Neighbor Algorithm as an Object Localization Technique

Using Passive RFID Tags

By Akshay Shetty

Thesis Director:

Professor Ivan Marsic

Technologies using identification by radio frequencies (RFID) are experiencing rapid development and healthcare is a major application area benefiting from it. Highly pervasive RFID enables remote identification, tracking and localization of the medical staff, patients, medications and equipment, thus increasing safety, optimizing in real-time management and providing support for new ambient-intelligent services. This thesis describes and evaluates an algorithm that enables object localization and tracking using passive RFID tags. This thesis also describes scenarios of how this technology can be used as a part of building a smart trauma resuscitation room by tracking the equipments. The main contribution of this thesis is the adaptation of the Weighted K-Nearest Neighbor Algorithm as a localization technique to track objects in a confined and crowded space by using passive RFID tags. The input parameter to the algorithm is the received signal strength indicator (RSSI), which gives a measure of back-scattered radio frequencies from passive tags. While using RFID technology special attention has to be given to the placement of antennas to get the optimum result. Therefore, we analyzed

various antenna placement configurations with mean error and error consistency as the two performance parameters. The detection of multiple tags and human occlusion are two major concerns while tracking tags in a confined space with many team members collaborating on solving a problem. The RF signal can be interrupted by people walking around randomly and holding multiple (tagged) instruments at the same time. While the algorithm worked fine when tracking multiple tags, we had to modify the experimental set-up and attach an antenna onto the ceiling (which we call a vertical antenna), so that even if all the wall antennas are blocked we get at least one input parameter to base our localization decision on. We evaluated the algorithm for different combinations of configurations and number of neighbors, and achieved the following results.

The best results were obtained for the 3 antennae (placed orthogonally) configuration considering the 4 nearest neighbors wherein a mean error rate of 15% of the maximum possible error was achieved under ideal conditions. We tested the algorithm for different human occlusion scenarios i.e. blocking 1 or 2 wall antennas, standing in random positions and then roaming in the field area randomly. The mean error rate for the standing scenario was measured as 20% of the maximum possible error and 18% in the case of roaming configuration. The error was found to be consistently within our defined maximum error for 100% of the recorded readings.

The results obtained were found to be satisfactory for our application where, more than the exact location of the object, knowing whether the object is within a particular region is good enough for the users to know what task is being carried out in the trauma bay. Also the algorithm holds good in an indoor environment having a lot of factors and materials which affect the RF signal disrupting accurate calculation of the location co-ordinates. The algorithm does not require extensive data collection prior to implementation which makes it easily deployable in any environment. Apart from the problems mentioned there are some other factors like materials on which the tags are attached and orientation of tags which were found to be potential hindrances for accurate localization. Acceptable solutions to these problems form a part of our future work.

Dedication

To my Parents

Acknowledgement

Quite evidently I would like to start off by thanking Professor Marsic for giving me the opportunity to work on a very interesting, practical and well-defined problem. I will forever be indebted to him for his constant support, frank opinions, directional guidance and his ability to make me visualize unimaginable problem situations. I am grateful to my labmate, Siddika Parlak for always being available for fruitful as well as silly discussions. I thank all my friends at Rutgers, especially Shivangi, Sneha and Vaidehi for literally taking care of me during my stay here at Rutgers. I would like to express my gratitude to Bipin uncle for being there when it mattered the most. A special mention for my brother, who has always been a means of unadulterated support and the comic relief in my life, cannot imagine life without you. I would like to dedicate this work to my late father who has always been a means of inspiration. And finally I would like to thank my mother, whose sheer grit and determination has got me to where I am today, I owe everything to you.

Contents

List of illustrations

Chapter 1

1. Introduction

1.1 Motivation

Automatic identification, or auto ID for short, is the broad term given to a host of technologies that are used to help machines identify objects [1]. Auto identification is often coupled with automatic data capture. That is, companies want to identify items, capture information about them and somehow get the data into a computer without having employees type it in. The aim of most auto-ID systems is to increase efficiency, reduce data entry errors and free up staff to perform more value-added functions, such as providing customer service. There is a host of technologies that fall under the auto-ID umbrella. These include bar codes, smart cards, voice recognition, some biometric technologies (retinal scans, for instance), optical character recognition (OCR) and radio frequency identification (RFID).

Radio frequency identification, or RFID, is a generic term for technologies that use radio waves to automatically identify people or objects. There are several methods of identification, but the most common is to store a serial number that identifies a person or object, and perhaps other information, on a microchip that is attached to an antenna (the chip and the antenna together are called an RFID transponder or an RFID tag). The antenna enables the chip to transmit the identification information to a reader. The reader converts the radio waves reflected back from the RFID tag into digital information that can then be passed on to computers that can make use of it.

RFID is not necessarily "better" than bar codes. The two are different technologies and have different applications, which sometimes overlap. The big difference between the two is bar codes are line-of-sight technology. That is, a scanner has to "see" the bar code to read it, which means people usually have to orient the bar code toward a scanner for it to be read. Radio frequency identification, by contrast, doesn't require line of sight. RFID tags can be read as long as they are within range of a reader. Bar codes have other shortcomings as well. If a label is ripped or soiled or has fallen off, there is no way to scan the item, and standard bar codes identify only the manufacturer and product, not the unique item. The bar code on one milk carton is the same as every other, making it impossible to identify which one might pass its expiration date first.

RFID is a proven technology that's been around since at least the 1970s. Up to now, it's been too expensive and too limited to be practical for many commercial applications. But if tags can be made cheaply enough, they can solve many of the problems associated with bar codes. Radio waves travel through most non-metallic materials, so they can be embedded in packaging or encased in protective plastic for weatherproofing and greater durability. And tags have microchips that can store a unique serial number for every product manufactured around the world.

Many companies have invested in RFID to get the advantages it offers. These investments are usually made in closed-loop systems—that is, when a company is tracking goods that never leave its own control. That's because some existing RFID systems use proprietary technology, which means that if company A puts an RFID tag on a product, it can't be read by Company B unless they both use the same RFID system from the same vendor. Another reason is the price. If a company tracks assets within its own four walls, it can reuse the tags over and over again, which is cost-effective. But for a system to work in an open supply chain, it has to be cheap because the company that puts the tag on a case or pallet is unlikely to be able to reuse it.

One issue for reluctance in use of RFID technology is standards. There are well-developed standards for low- and high-frequency RFID systems, but most companies want to use UHF in the supply chain because it offers longer read range—up to 20 feet under good conditions. UHF technology is relatively new, and standards weren't established until recently. Another issue is cost. RFID readers typically cost $1,000 or more. Companies would need thousands of readers to cover all their factories, warehouses and stores. RFID tags are also fairly expensive—20 cents or more—which makes them impractical for identifying millions of items that cost only a few dollars.

RFID systems can be used just about anywhere, from clothing tags to missiles to pet tags to food - anywhere that a unique identification system is needed. The tag can carry information as simple as a pet owners name and address or the cleaning instruction on a sweater to as complex as instructions on how to assemble a car.

Here are a few examples of how RFID technology is being used in everyday places [1]:

- RFID systems are being used in some hospitals to track a patient's location, and to provide real-time tracking of the location of doctors and nurses in the hospital. In addition, the system can be used to track the whereabouts of expensive and critical equipment, and even to control access to drugs, pediatrics, and other areas of the hospital that are considered "restricted access" areas.
- RFID chips for animals are extremely small devices injected via syringe under skin. Under a government initiative to control rabies, all Portuguese dogs must be

RFID tagged by 2007. When scanned the tag can provide information relevant to the dog's history and its owner's information.

- RFID in retail stores offer real-time inventory tracking that allows companies to monitor and control inventory supply at all times.
- The Orlando/Orange County Expressway Authority (OOCEA) is using an RFID based traffic-monitoring system, which uses roadside RFID readers to collect signals from transponders that are installed in about 1 million E-Pass and SunPass customer vehicles.

Apart from the above mentioned applications RFID technology is fast emerging as a solution to building smart hospitals and our thesis covers a small aspect of this particular area of application. Currently RFID is being used in the following healthcare applications

Patient Identification

Many health professionals are concerned about the growing number of patients who are misidentified before, during or after medical treatment. Indeed, patient identification error may lead to improper dosage of medication to patient, as well as having invasive procedure done. Other related patient identification errors could lead to inaccurate lab work and results reported for the wrong person, having effects such as misdiagnoses and serious medication errors [2].

In order to cut these clinical errors, to improve patient care and security and also to improve administration and productivity, several RFID-based patient identification and tracking pilot projects have been launched during the last two years.

Blood Tracking

Mis-transfusion errors (i.e. blood transfusion of the incorrect type or blood given to the wrong patient) are unacceptably frequent and serious in hospitals. mis-transfusions typically result from an error made during the bedside check just prior to transfusion. Studies[2] have documented that such errors are most likely to occur among surgical patients. Currently the bedside check is done by humans using eye-readable information, and in operating rooms this task is particularly difficult. Indeed, blood is often given under circumstances of extreme urgency and distraction. Patients are unconscious during the transfusion and cannot state their name, and caregivers in the operating rooms may not "know" the patient as well as nurses on non-surgical floors.

To address the issue of the bedside transfusion check one should take advantage of new technology. Two machine-readable technologies are candidates for the automation of these checks: bar code technology and RFID. Barcodes are unsuitable for bedside checks because they require line-of-sight so that a handheld laser can read a flat surface with the code. This constraint represents an important practical obstacle, especially in operating rooms where the patient is covered with surgical drapes.

Smart Operating Theatres

In the smart hospital the patients get a RFID-tagged wristband containing relevant information and a digital picture of them. The photograph allows the clinical team to easily confirm they have the right patient, and the electronic record ensures they perform the correct procedure. If the wrong patients enter the operating room, the medical staff is automatically and instantly warned of the mismatch. Thus radio tagging makes the operating theatre safer and more efficient. Moreover the risk of litigation resulting from surgery mistakes and the costs they generate should be significantly reduced [2].

Our application

Equipment Tracking in the Trauma bay

Trauma is a leading cause of death and disability in children and young adults. Because care after injury has an important impact on outcome, initial management of the injured patient should be rapid, efficient and error-free. The initial resuscitation of an injured patient is prone to errors because workers often manage patients with unstable or fluctuating medical status, whose detailed medical history is commonly not available and who require time critical decisions. Medical errors are prevented mainly through provider experience and redundancies in the evaluation process. The potential errors increase when treating injured children, because variable anatomy and physiology at different areas prevent standardization.

Information exchange during trauma resuscitation is unique, since clinical data is simultaneously and rapidly accrued by several providers and used to make decisions within minutes. The potential for errors increases because providers need to focus on their tasks while maintaining constant awareness of tasks of others. While information technologies have been used to support highly strained cognitive activities in time-critical settings, these methodologies have not been applied for trauma evaluation. Previous studies have focused primarily on individual team members who are operating sensors or computerized controls. This type of analysis will not be effective in environments such as trauma resuscitation with structured teams of members with different roles, teamwork being highly interdependent, and members relying primarily on judgment and experience rather than sensors. Studying teamwork in this environment therefore cannot rely on automatic data capture.

The resuscitation bay in hospitals is a very dynamic and busy room when used. An injured patient enters the room and the room becomes crowded by doctors, nurses, radiologists and other personnel. To the unfamiliar eye this situation seems rather chaotic. But as time progresses and less people are moving around a smaller group of medical personnel remains in the room. For all personnel in this room the situation itself is quite familiar and less chaotic then seen by a new observer. During resuscitation every person in the room has a predetermined role. The composition of teams is usually very similar at different facilities and includes an attending surgeon, surgical residents, an anesthesiologist, an orthopedic surgeon, nurses, a respiratory therapist, a pharmacist and an x-ray technician. Not everyone is needed during the complete resuscitation and people arrive and leave the resuscitation bay at regular intervals. It is essential for the group working on the patient to combine their efforts into an orchestrated process with the goal to stabilize the patient.

Part of the resuscitation bay is a decision support system able to devise a plan for resuscitating the patient and advising or critiquing a surgeon which can minimize errors. By supplying the expert system with data about the patient and its environment, the system can reason about resuscitation plans to stabilize the patient. Several sources of information are available in the resuscitation bay to supply the expert system with useful knowledge. The expert system preferably wants to critique a plan that is being executed. Then information about the tasks the resuscitation team is currently performing is very important [18]. This is where our system comes into play. By localizing the various equipments in the trauma bay, the objects that are in use can be identified giving the expert system an idea about the tasks being performed.

1.2 Research Questions

Since this an explored area of research in terms of the application, we were faced with many questions. We tried to answer some of the fundamental questions that would create an application area that can lead to further research. Some of the questions are listed below.

1. Scalability:

   Whether the system can be easily deployed over a large area? We will be deploying the system over an approximate area of 15x15 feet.

2. Accuracy:

   How accurate can the system be? In trauma care accuracy is of prime importance as it is a very sensitive and life saving application. We are looking at a system which does not result in an error rate of more than 20% of the maximum possible error.

3. Cost:

   How much will the entire system cost? The expense of implementing the entire system should not be exorbitant. The cost of the system should be in the order of a few thousands of dollars.

4. Performance:

   How fast will the system be in tracking objects? A tracking application especially in a trauma-care environment should be pseudo-real time if not real-time. The system should be fast enough to cater to average human walking speeds i.e. 4.5mph.

5. Multiple tag detection:

   Can multiple tags be detected? There can be multiple equipments present in the field area at a particular time. The system should be able to track around 10 equipments.

6. Overcoming Human Occlusion:

   Can the system provide satisfactory results under human interruption? The area of application will have doctors and nurses moving around randomly and the system should not result in an error rate of more than 20 % of the maximum possible error as mentioned earlier under such conditions.

In this thesis we present a system that answers these questions. We also present the analysis of the findings that show some clear areas of improvements.

2. Background

2.1 How does RFID work?

A Radio-Frequency IDentification system has three parts [3]:

- A scanning antenna
- A transceiver with a decoder to interpret the data
- A transponder - the RFID tag - that has been programmed with information.

The RFID tag may be of one of two types. Active RFID tags have their own power source; the advantage of these tags is that the reader can be much farther away and still get the signal. Even though some of these devices are built to have up to a 10 year life span, they have limited life spans. Passive RFID tags, however, do not require batteries, and can be much smaller and have a virtually unlimited life span.

The scanning antenna puts out radio-frequency signals in a relatively short range. The RF radiation does two things:

- It provides a means of communicating with the transponder (the RFID tag) AND
- It provides the RFID tag with the energy to communicate (in the case of passive RFID tags).

This is an absolutely key part of the technology; RFID tags do not need to contain batteries, and can therefore remain usable for very long periods of time (maybe decades).

The scanning antennas can be permanently affixed to a surface; handheld antennas are also available. They can take whatever shape you need; for example, you could build them into a door frame to accept data from persons or objects passing through.

When an RFID tag passes through the field of the scanning antenna, it detects the activation signal from the antenna. That "wakes up" the RFID chip, and it transmits the information on its microchip to be picked up by the scanning antenna.

RFID tags can be read in a wide variety of circumstances, where barcodes or other optically read technologies are useless.

- The tag need not be on the surface of the object (and is therefore not subject to wear)
- The read time is typically less than 100 milliseconds
- Large numbers of tags can be read at once rather than item by item.

An RFID tag is an active tag when it is equipped with a battery that can be used as a partial or complete source of power for the tag's circuitry and antenna. Some active tags contain replaceable batteries for years of use; others are sealed units. (Note that It is also possible to connect the tag to an external power source.)

The major advantages of an active rfid tag are [4]:

- It can be read at distances of one hundred feet or more, greatly improving the utility of the device
- It may have other sensors that can use electricity for power.

The problems and disadvantages of an active RFID tag are:

- The tag cannot function without battery power, which limits the lifetime of the tag.
- The tag is typically more expensive, often costing $20 or more each

- The tag is physically larger, which may limit applications.
- The long-term maintenance costs for an active RFID tag can be greater than those of a passive tag if the batteries are replaced.
- Battery outages in an active tag can result in expensive misreads.

A passive tag is an RFID tag that does not contain a battery; the power is supplied by the reader. When radio waves from the reader are encountered by a passive rfid tag, the coiled antenna within the tag forms a magnetic field. The tag draws power from it, energizing the circuits in the tag. The tag then sends the information encoded in the tag's memory.

The major disadvantages of a passive rfid tag are:

- The tag can be read only at very short distances, typically a few feet at most. This greatly limits the device for certain applications.
- It may not be possible to include sensors that can use electricity for power.
- The tag remains readable for a very long time, even after the product to which the tag is attached has been sold and is no longer being tracked.

The advantages of a passive tag are:

- The tag functions without a battery; these tags have a useful life of twenty years or more.
- The tag is typically much less expensive to manufacture.
- The tag is much smaller (some tags are the size of a grain of rice). These tags have almost unlimited applications in consumer goods and other areas.

Since the surgical instruments in hospitals are small, it made sense for us to go in for a system using passive rfid tags because they are relatively small as compared to active tags. Also active tags are bulky as they carry a power source which makes it inapplicable for our application.

2.2 Approaches to Positioning

Radio propagation in indoor environment is subject to numerous problems such as severe multipath, rare line-of sight (LOS) path, absorption, diffraction, and reflection. Since signal cannot be measured very precisely, several indoor localization algorithms have been proposed in the literature. They can be classified in three families: distance estimation, scene analysis, and proximity [5].

2.2.1 Distance Estimation

This family of algorithms uses properties of triangles to estimate the target's location. The triangulation approach consists in measuring the angle of incidence (or Angle Of Arrival - AOA) of at least two reference points. The estimated position corresponds to the intersection of the lines defined by the angles. On the contrary, the lateration approach, estimates the position of the target by evaluating its distances from at least three reference points. The range measurements techniques use Received Signal Strength (RSS), Time Of Arrival (TOA), Time Difference Of Arrival (TDOA), or Received Signal Phase (RSP).

1) RSS: The attenuation of emitted signal strength is function of the distance between the emitter and the receiver. The target can thus be localized with at least three reference points and the corresponding signal path losses due to propagation. Several empirical and theoretical models have been proposed to translate the difference between the transmitted and the received signal strength into distance estimation. The RSS based systems usually need on-site adaptation in order to reduce the severe effects of multipath fading and shadowing in indoor environments [7].

2) TOA: The distance between a reference point and the target is also proportional to the propagation time of signal. TOA-based systems need at least three different measuring units to perform a lateration for 2-D positioning. However, they also require that all transmitters and receivers are precisely synchronized and that the transmitting signals include timestamps in order to accurately evaluate the traveled distances. If more than three reference points are available, the least-squares algorithm or one of its variants can be used in order to minimize the localization error [8].

3) TDOA: The principle of TDOA lies on the idea of determining the relative location of a targeted transmitter by using the difference in time at which the signal emitted by a target arrives at multiple measuring units. Three fixed receivers give two TDOAs and thus provide an intersection point that is the estimated location of the target. This method requires a precise time reference between the measuring units. Like TOA, TDOA has other drawbacks. In indoor environments, a LOS channel is rarely available. Moreover, radio propagation often suffers from multipath effects thus affecting the time of flight of the signals [9].

4) RSP: The RSP method, also called Phase Of Arrival (POA), uses the delay, expressed as a fraction of the signal's wavelength, to estimate distance. It requires transmitters placed at particular locations and assumes that they emit pure sinusoidal signals. The localization can be performed using phase measurements and the same algorithm than TOA or phase difference measurements and the same algorithm than TDOA. The disadvantage of the RSP method when applied in indoor environments is that it strongly needs a LOS signal path to limit localization errors.

5) AOA: AOA consists in calculating the intersection of several direction lines, each originating from a beacon station or from the target. At least two angles, measured with directional antennae or with an array of antennae and converted in direction lines, are needed to find the 2-D location of a target. Nevertheless, this technique requires complex and expensive equipments and notably suffers from shadowing and multipath reflections [10].

Example – Trilateration

Trilateration is a method for determining the intersections of three sphere surfaces given the centers and radii of the three spheres.

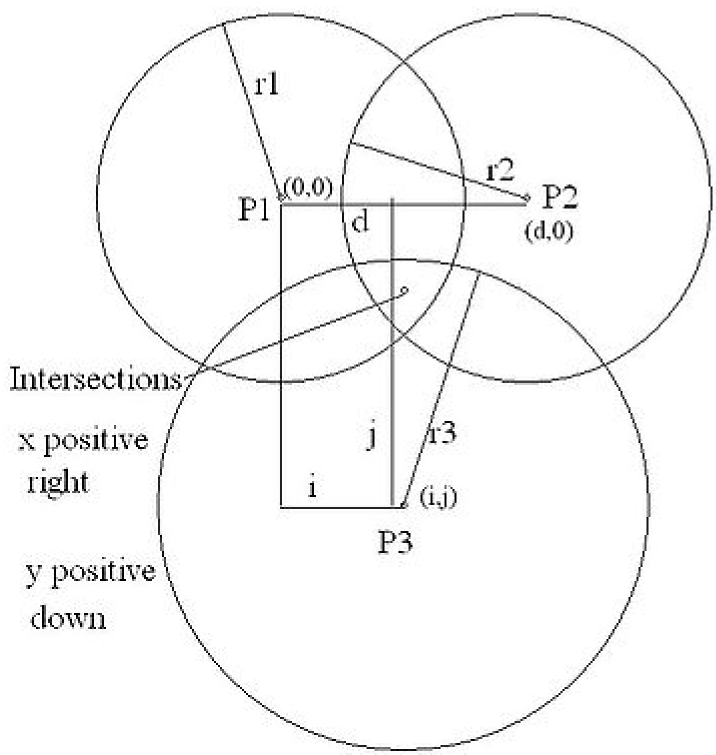

Figure 1: Geometry for Trilateration
Source : www.absoluteastronomy.com

A mathematical derivation for the solution of a three-dimensional trilateration problem can be found by taking the formulae for three spheres and setting them equal to each other. To simplify the calculations, we apply three constraints to the centers of these spheres; we assume all three spheres are centered on the z=0 plane, one is at the origin, and one other is on the x-axis. It is possible to transform any set of three points to comply with these constraints, find the solution point, and then reverse the translation to find the solution point in the original coordinate system. The formulae for the 3 spheres are as follows

$r1^2 = x^2 + y^2 + z^2$

$r2^2 = (x-d)^2 + y^2 + z^2$

$r3^2 = (x-i)^2 + (y-i)^2 + z^2$ where, (x,y,z) is the location of tag.

Solving the above equations for x, y, and z, the location of the tag can be determined.

Although this method is very simple to implement using the range measurement techniques directly to estimate the tag's location does not yield required results.

2.2.2 Scene analysis

Scenes analysis approaches are composed of two distinctive steps. First, information concerning the environment (fingerprints) is collected. Then, the target's location is estimated by matching online measurements with the appropriate set of fingerprints. Generally, RSS-based fingerprinting is used. The two main fingerprinting-based techniques are: k-nearest neighbor (kNN) also known as radio map, and probabilistic methods.

kNN - The kNN method consists in a first time in measuring RSS at known locations in order to build a database of RSS that is called a radio map. Then, during the online phase, RSS measurements linked to the target are performed to search for the *k* closest matches in the signal space previously-built. Root mean square errors principle is finally applied on the selected neighbors to find out an estimated location for the target [11] [12] [13].

Probabilistic Approach - The problem stated in probabilistic approaches is to find the location of a target assuming that there are *n* possible locations and one observed signal strength vector during the online phase according to posteriori probability and Bayes formula. Thus, the location with the highest probability is chosen. Generally, probabilistic methods involve different stages such as calibration, active learning, error estimation, and tracking with history [14].

Example : Kalman Filtering

This approach utilizes reference tags. The first step consists in calculating with RSS measurements from 2 readers the distance D_i between each reference tag and the target tag. The location of the target tag is obtained by solving with the minimum mean squared error algorithm, the system of non-linear equations.

$(x_i - x_e)^2 + (y_i - y_e)^2 = {D_i}^2$ for i = 1, 2, …….n

The second step consists in building a probabilistic map of the error measurement for the reader's detection area. The first step is applied for each reference tag in order to calculate their corresponding error probability distribution function with the help of their estimated location and real location. The Kalman filter is then used iteratively in this online map to reduce the effect of RSS error measurement and thus to improve the accuracy of the localization.

But this particular method requires extensive data collection to be done and also takes up a lot of time in processing and analyzing data which makes it a little difficult to be deployed in real time.

2.2.3 Proximity

The last type of localization techniques in indoor environments is based on proximity. This approach relies on dense deployment of antennae. When the target enters in the radio range of a single antenna, its location is assumed to be the same as this receiver. When more than one antenna detects the target, the target is assumed to be collocated with the one that receives the strongest signal. Again this approach is very basic and easy to implement. However, the accuracy is on the order of the size of the cells.

Example : Concentric Power Levels

One of the popular methods of estimating distances is by controlling the power levels of the antenna and getting a list of tags read at each level. The power levels can be then correlated to a distance measure and the tags can be positioned. Though this method can be used on RFID readers that do not supply the RSSI information and would make it a more generalized solution, we found this method was not any more effective than using the RSSI parameter for estimation. Also in order to estimate the entire set of tags in the read range we need to cycle through at least a few power levels depending on the granularity of the distance estimation. Also tags which are close enough would be found in multiple power levels hence might affect the read counts of all the tags.

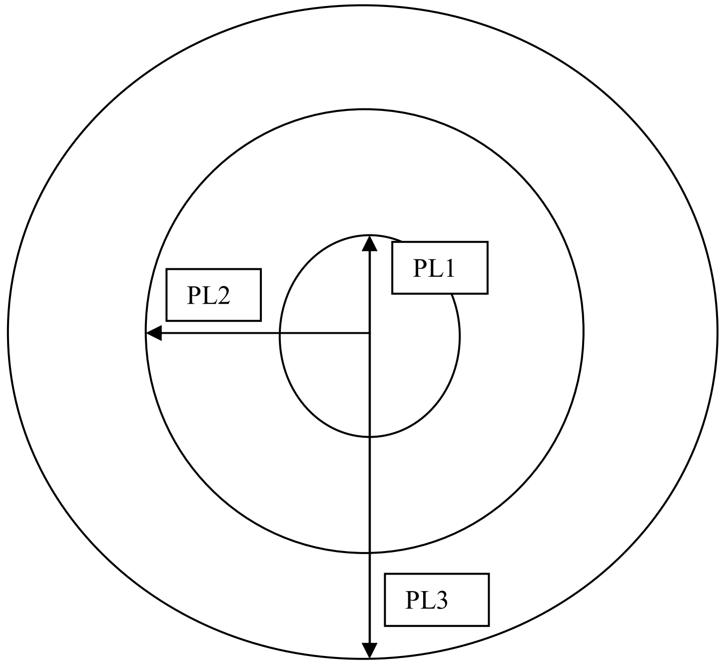

Figure 2: Concentric circles formed by increasing power levels

For our system we will be using a hybrid approach wherein we use the kNN algorithm to find the nearest reference positions and perform lateration on their co-ordinates to get the unknown location of tracking object.

3. Proposed Approach (Weighted K-Nearest Neighbor)

3.1 WKNN [16] Algorithm

The WKNN algorithm has been implemented using active RFID tags in the LANDMARC technique but in our work we use it to track passive tags and discuss its feasibility. We have also tried various combinations of distance between reference tags and antenna positions to get optimum results. We believe WKNN will yield results close to our expectations defined earlier as it does not localize the tag directly based on one of the range measurement techniques and rather uses them as a parameter to be worked on with an algorithm. Also it does not require extensive pre-processing of data as compared to some other techniques which makes it easier to employ as well.

As seen in the flow chart in figure 3, implementation of the algorithm consists of some pre-processing of data before we calculate the exact position of tag in real time. We place the tag at certain pre-determined reference positions and register the RSSI values received from each antenna at that particular position in a database called as the radio map. So, in the radio map we have a corresponding RSSI value registered for the tag for each reference position.

While localizing the tag in real time, an average of all collected RSSI values for a period of 1.5 seconds is taken as an input parameter from each antenna. The reason for using 1.5 seconds as the data collection time span is justified in section 5.6. The RSSI values are then compared with the RSSI values of all the reference positions using the Euclidean distance formula resulting in what we call as the relativity parameter ‘E’. The ‘K’ nearest neighbors (reference positions) are determined as the ones which have the lowest value of ‘E’.

The actual co-ordinates of the tracking tag can be determined by applying a lateration formula on the co-ordinates of the ‘K’ nearest reference positions.

3.2 WKNN Flow-Chart

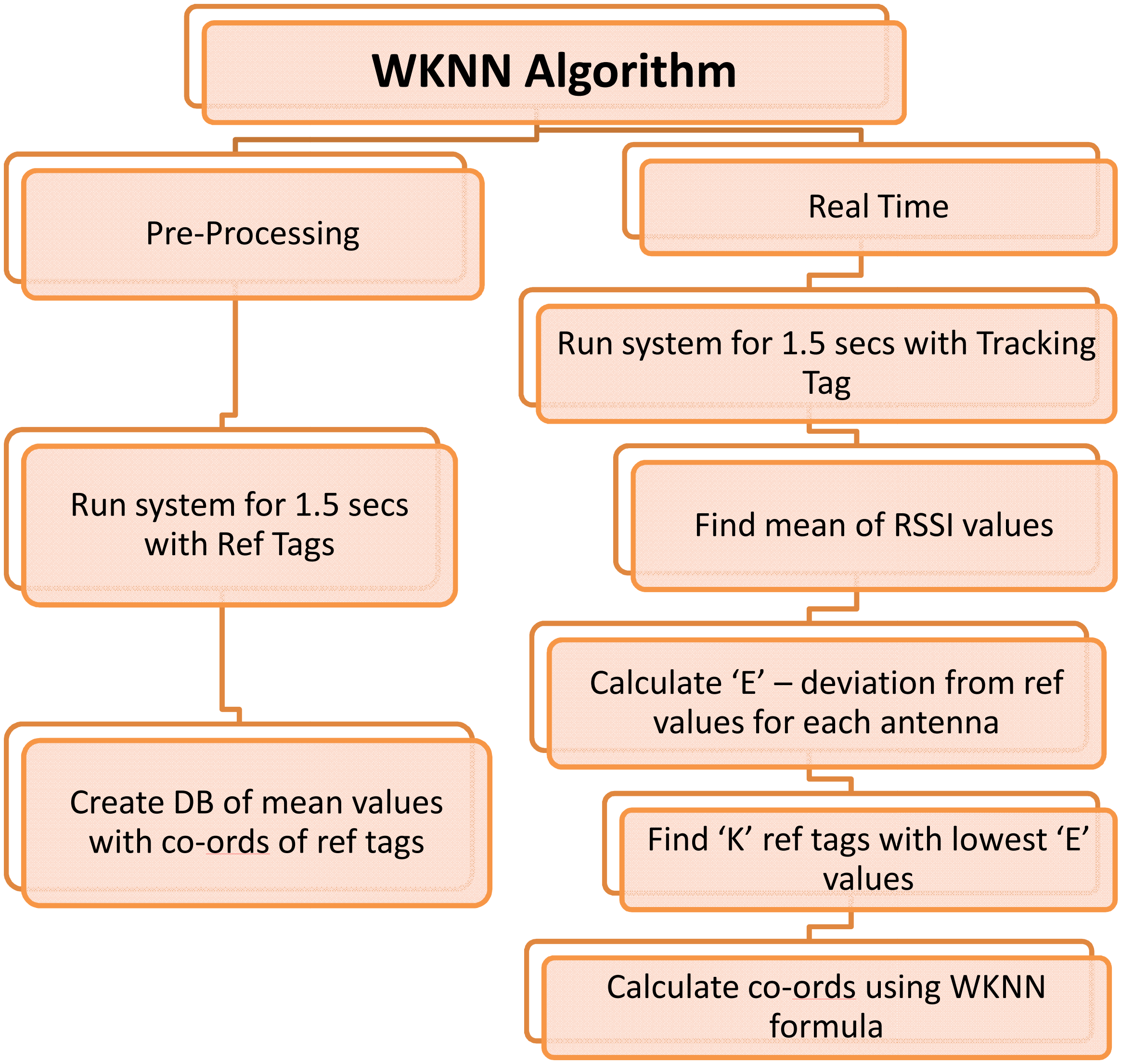

Figure 3: WKNN Algorithm flow-chart

3.3 Mathematical Interpretation:

N = Number of readers;

R = Reference positions' measured signal strength vector;

S = Object tag's measured signal strength vector;

E = Measure of closeness (Eucledian distance) between reference position vector (R) and object tag vector (S).

(X, Y) = Co-ordinates of reference position.

(x, y) = Co-ordinates of object tag's position.

$$E^{(m)} = \left(\sum_{i=1}^{N}(R_i^{(m)} - S_i)^2\right)^{1/2}$$

$$(x, y) = \sum_{i=1}^{k}\left\{\frac{\left(\frac{1}{E_i^2}\right).(X_i,Y_i)}{\sum_{j=1}^{k}\left(\frac{1}{E_j^2}\right)}\right\}$$

3.4 General Interpretation with example:

Let us take an example wherein we calculate the position of the object tag by considering 4 nearest neighbors i.e. K = 4.

Procedure:

1) Measure the RSSI signal vector (RSSI values from 3 antennas) for the object tag – R_1, R_2, R_3.

2) Find the proximity of these values to all the reference position values i.e. if we have 12 reference positions, we calculate $E^{(1)}$, $E^{(2)}$ … $E^{(12)}$ using the following Eucledian distance formula.

$E^{(1)} = ((R_1^{(1)} - S_1)^2 + (R_2^{(1)} - S_2)^2 + (R_3^{(1)} - S_3)^2)^{1/2}$

$E^{(2)} = ((R_1^{(2)} - S_1)^2 + (R_2^{(2)} - S_2)^2 + (R_3^{(2)} - S_3)^2)^{1/2}$

.

.

.

$E^{(m)} = ((R_1^{(m)} - S_1)^2 + (R_2^{(m)} - S_2)^2 + (R_3^{(m)} - S_3)^2)^{1/2}$

3) Find the reference positions with 4 lowest 'E' values and term them as E_1, E_2, E_3 & E_4.

4) Let the co-ordinates of these reference positions be (X_1, Y_1), (X_2, Y_2), (X_3, Y_3) &

(X_4, Y_4).

5) Calculate the unknown co-ordinates (x, y) of the object tag using

$$(x, y) = \frac{\left(\frac{1}{E_1^2}\right)}{E^2}.(X_1, Y_1) + \frac{\left(\frac{1}{E_2^2}\right)}{E^2}.(X_2, Y_2) + \frac{\left(\frac{1}{E_3^2}\right)}{E^2}.(X_3, Y_3) + \frac{\left(\frac{1}{E_4^2}\right)}{E^2}.(X_4, Y_4)$$

Where $E^2 = \frac{1}{E_1^2} + \frac{1}{E_2^2} + \frac{1}{E_3^2} + \frac{1}{E_4^2}$

3.5 Testing Parameters

Mean Error

The error in localizing the object tag is defined as the point to point distance between the expected tag co-ordinates and the predicted tag co-ordinates. We calculate this distance using the following Eucledian distance formula.

Suppose (x_1, y_1) are the expected co-ordinates and (x_2, y_2) are the predicted co-ordinates, then the error is given by

$$\text{Error} = \sqrt[2]{((x_1 - x_2)^2 + (y_1 - y_2)^2)}$$

Percentage of Maximum Error

After calculating the mean error, another interesting testing parameter that we consider for our experiments is mean error as a percentage of the maximum possible error. When the algorithm is applied to a particular defined area, the mean error by itself does not indicate the accuracy of the algorithm as a function of the field area which is a very important factor while localization. Therefore, by considering the maximum possible error as a comparison factor we have tried to measure the accuracy as a function of the applied field area as well.

Histogram Analysis

Apart from the above mentioned parameters where accuracy is measured quantitatively, consistency in the error is also a major factor in deciding the feasibility of an algorithm. For this reason we study the histogram of all obtained results and calculate the percentage of those experiments which yield acceptable mean error values.

3.6 Testing procedure

The feasibility of the algorithm is based on 3 fundamental factors i.e. the mean error, mean error as a percentage of maximum possible error and consistency of this mean error lying within a certain specified maximum threshold. To test these factors we define 2 set-ups. The first set-up is where we implement the algorithm on a small area for example a surgical table and after testing for various test cases, we implement the algorithm onto a bigger set-up i.e. a 9x9 feet field area.

Set-up 1 (Surgical table)

In the first set-up, the testing table is divided into 6 15x15 inch blocks labeled 11, 12, 21, 22, 31 and 32 named after 3 rows and 2 columns. For every configuration discussed in section 4.2, 10 experiments are performed in each block and the error is calculated. The error of any configuration is the mean error of 60 experiments over all these 6 blocks. Many different test cases were applied onto this set-up so that we had a fair idea about the characteristics of the algorithm before implementing it on a larger area.

Set-up 2 (9x9 feet field area)

The second set-up involves implementation of the algorithm onto a larger area i.e. 9x9 feet after testing it for various configurations on a surgical table much smaller in field area. The area is divided into 9 3x3 feet testing blocks named as 11, 12, 13, 21, 22, 23, 31, 32 and 33 after 3 rows and 3 columns. Again 10 experiments are performed in each testing blocks for 9 center test points and 12 edge test points.

4. Implementation

4.1 Hardware

We employed RFID Readers, Antennae and Tags manufactured by Alien, the specifications for which have been tabulated below.

RFID Reader

Name	Alien Multi-Port General Purpose RFID Reader
Model Number	ALR 9900
Architecture	Point-to-multipoint reader network, mono-static antenna
Operating Frequency	902.75 MHz – 927.25 MHz
Hopping Channels	50
Channel Spacing	500 KHz
Channel Dwell Time	< 0.4 seconds
RF Transmitter	< 30 dBm at the end of 6 m LMR-195 cable.
Modulation Method	Phase Reversal – Amplitude Shift Keying (PR-ASK)
20 db Modulation Bandwidth	< 100 KHz
RF Receiver	2 Channels
Power Consumption	30 Watts
Communications Interface	RS-232 (DB-9 F), TCPI/IP (RJ-45)
Inputs/Outputs	4 coax antenna, 4 inputs/8 outputs (optically isolated), RS-232 com port, LAN, power
Dimensions	8 " (20.3 cm) x 7 " (17.8 cm) x 1.6 " (4.1 cm)
Weight	Approximately 1 kg (2.2 lbs)
Operating Temperature	-20°C to +50°C (-4 °F to +122°F)
LED Indicators	Power, Link, Active, Ant0-3, CPU, Read, Sniff, Fault (red)
Software Support	APIs, sample code, executable demo app (Alien Gateway)
Protocol Support	Comply with EPC Class 1 Gen 2 and 18000 – 6C
Compliance Certifications	FCC Part 15; FCCID: P65ALR9900 IOC: 4370A-ALR9900
Safety Certifications	cTUVus UL: 60950-1:2004 CAN/CSA: C22.2 No.60950-1-03

Figure 4: RFID Reader Specifications

Figure 5: Alien ALR 9900 RFID Reader

RFID Reader External Circular Polarized Antenna

Model	ALR-9611-CR and ALR-9611-CL
3 dB Beamwidth	E-plane: 65 degrees • H-plane: 65 degrees
Frequency	902-928 MHz
Gain (dBi)	6.0 dBiL (maximum)
Polarization	Circular
RF Connector	6 m LMR-195 with Reverse-Polarity TNC
VSWR	1.5:1
Dimensions	(cm) 22 x 27 x 4 • (in) 8.5 x 10.5 x 1.65
Weight	.57 kg • 1.25 lb

Figure 6: RFID Antenna Specifications

Figure 7: Alien ALR-9611 CR Antenna

RFID Tags

ISO/IEC 18000-6C	
EPCglobal Class 1 Gen 2	
Integrated Circuit	Alien Higgs-3
EPCglobal Certificate	950110126000001084
Operating Frequency	840–960 MHz
EPC Size	96 - 480 Bits
User Memory	512 Bits
TID	32 Bits
Unique TID	64 Bits
Access Password	32 Bits
Kill Password	32 Bits

Figure 8: Passive RFID tag Specifications

Figure 9: The squiggle passive RFID tag

4.2 Experimental Set-Up

Placement of antennas is a very important aspect of localization since we need to address the problem of scalability i.e. the entire field area is supposed to be covered by the antennas.

Since our application demands tracking of surgical instruments mainly during surgeries, we track the unknown tags by placing them over a 50in x 30in table. We started off by placing 2 antennas orthogonal to each other as seen in the figure 10 below. In all configurations the antenna was placed 50 inches above the floor level.

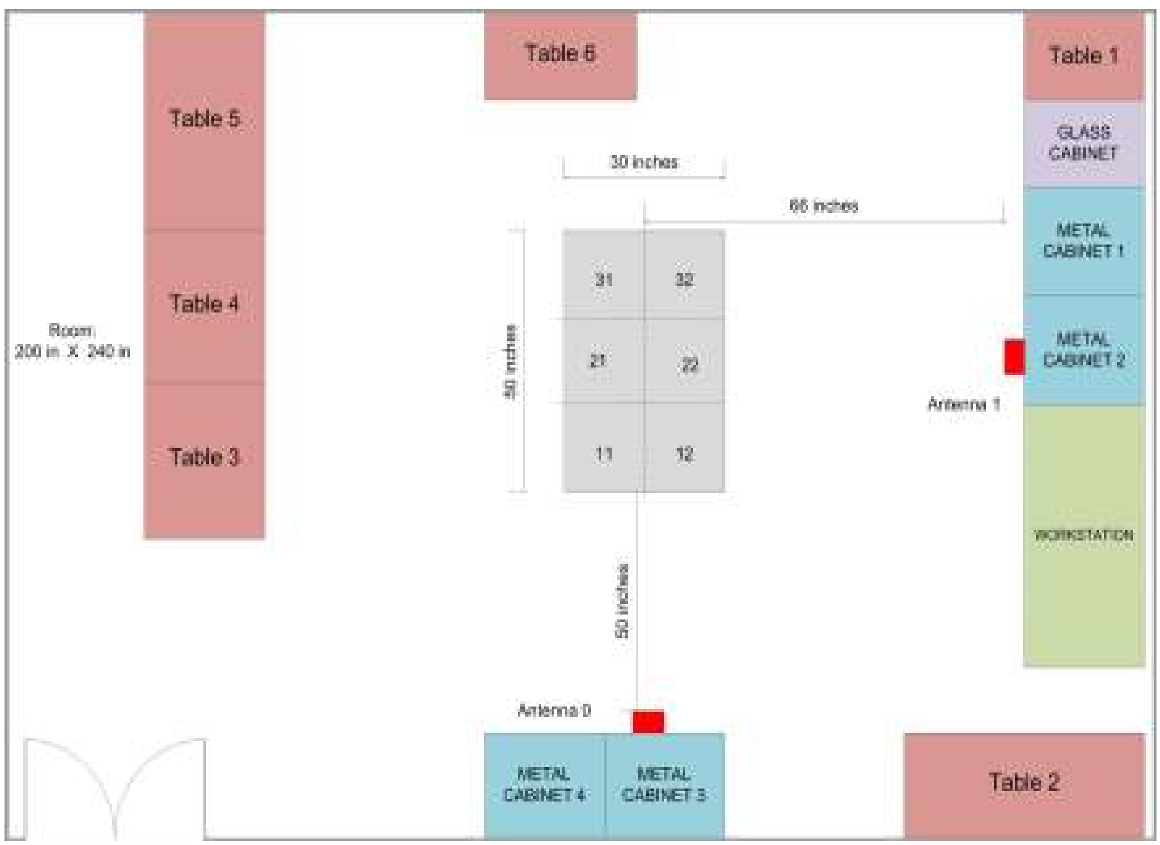

Figure 10: Floor plan for 2 orthogonally placed antennas

After analyzing the experimental results we realized that the output was way off the mark and probably using another antenna would enhance the result. This resulted in us trying out a different configuration, this time with three antennas as shown in the figure 11 below.

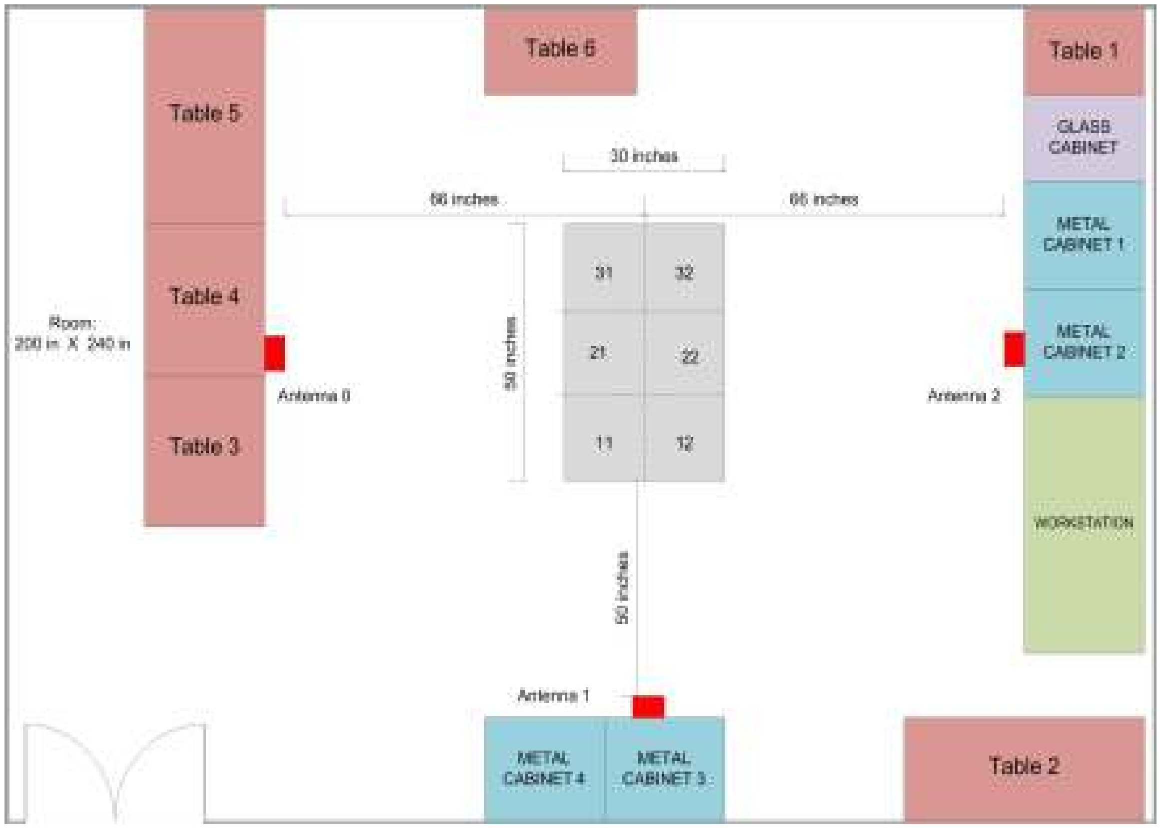

Figure 11: Floor plan for 3 wall antennas

The above mentioned configuration yielded satisfactory results but it was evident through experiments that it would not solve a very important problem of modern day localization i.e. human occlusion. A solution for efficient localization by eliminating effects of human occlusion was of prime importance especially in our application where moving human bodies within the field area is to be taken for granted.

This led to us employing a vertical antenna as the third antenna in addition to 2 wall antennas instead of 3 wall antennas. The vertical antenna makes sure that even if all the wall antennas are occluded we at least have one input parameter on which we can base our localization decision. For this set up as well we have tried 2 different configurations, 2 symmetrically positioned antennas with a vertical antenna and 2 orthogonally positioned antennas with a vertical antenna as shown in the figures 12 & 13 below.

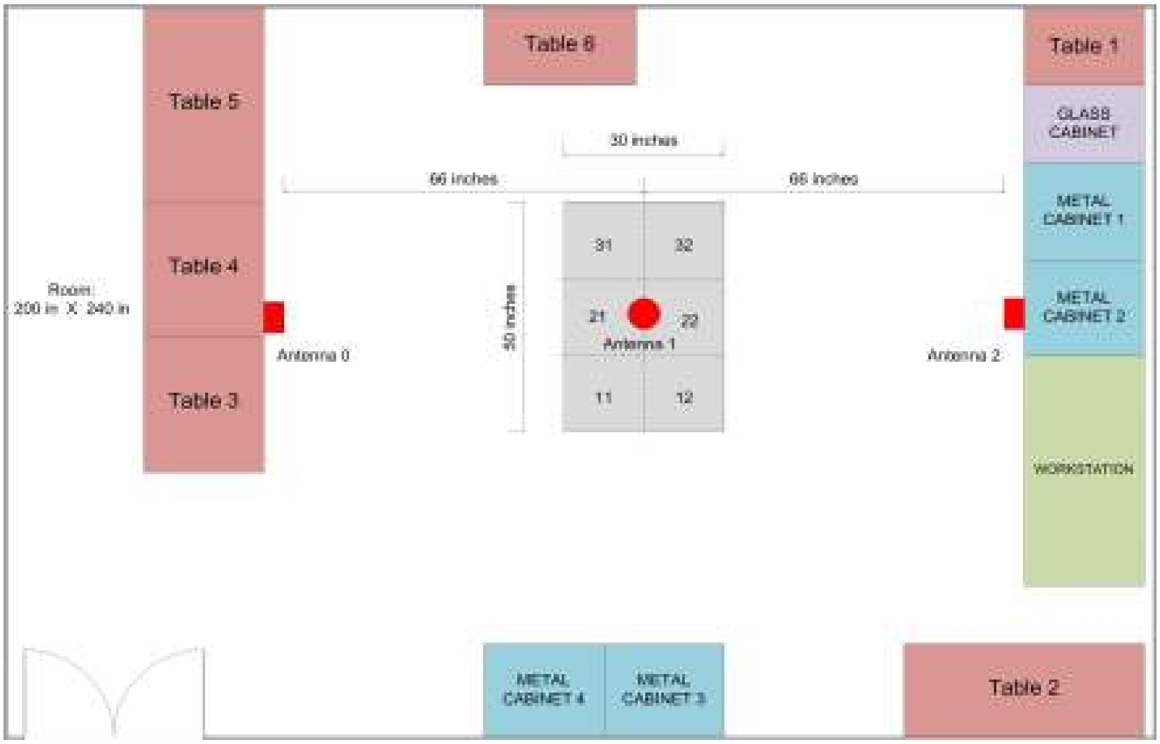

Figure 12: Floor plan for 2 symmetrically placed wall antennas and 1 vertical antenna

Figure 13: Floor plan for 2 orthogonally placed wall antennas and 1 vertical antenna

After testing the algorithm on the 50 x 30in table, we decided to test it on a larger area. We moved the table throughout a 9 x 9 feet area divided into 9 3 x 3 feet blocks i.e 16 reference positions, thereby recording the readings for the reference positions and also collecting readings for 21 test positions (9 center and 12 edge tag positions) with the orthogonal configuration as shown in figure 14.

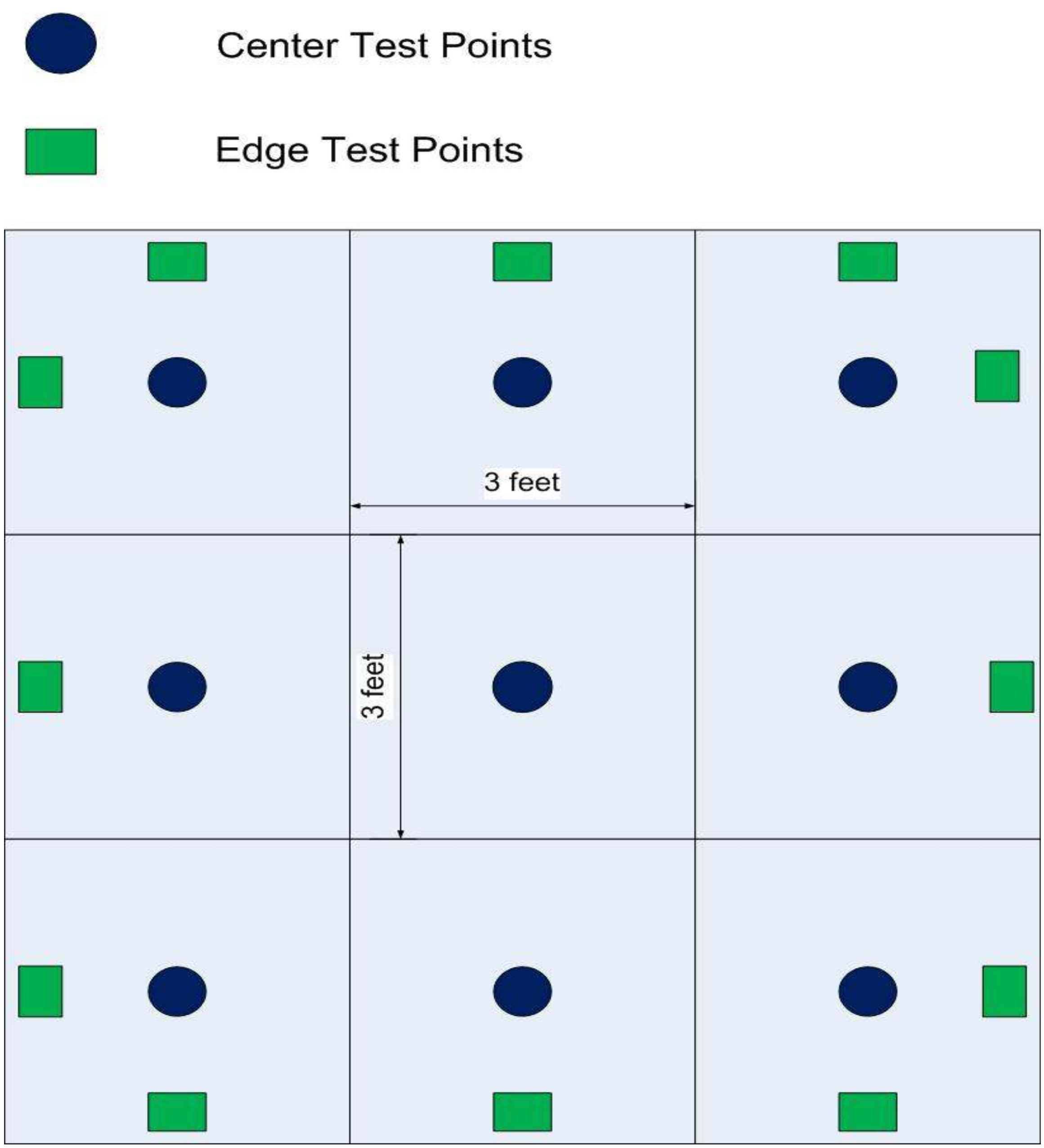

Figure 14: Floor plan for 9 x 9 feet field area under orthogonal configuration

The results discussed in Chapter 6 explain the use of all these configurations in detail.

5. Characteristics of Passive RFID Tags

Before starting off with setting up the experimental environment and testing the algorithm, it was extremely important for us to understand the behavior of the hardware components. As our algorithm uses RSSI values as a sole parameter for localization decisions we performed a few below mentioned tests to take note of the nature of registered RSSI values under certain specific circumstances.

5.1 Effect of distance of tag from reader on RSSI values.

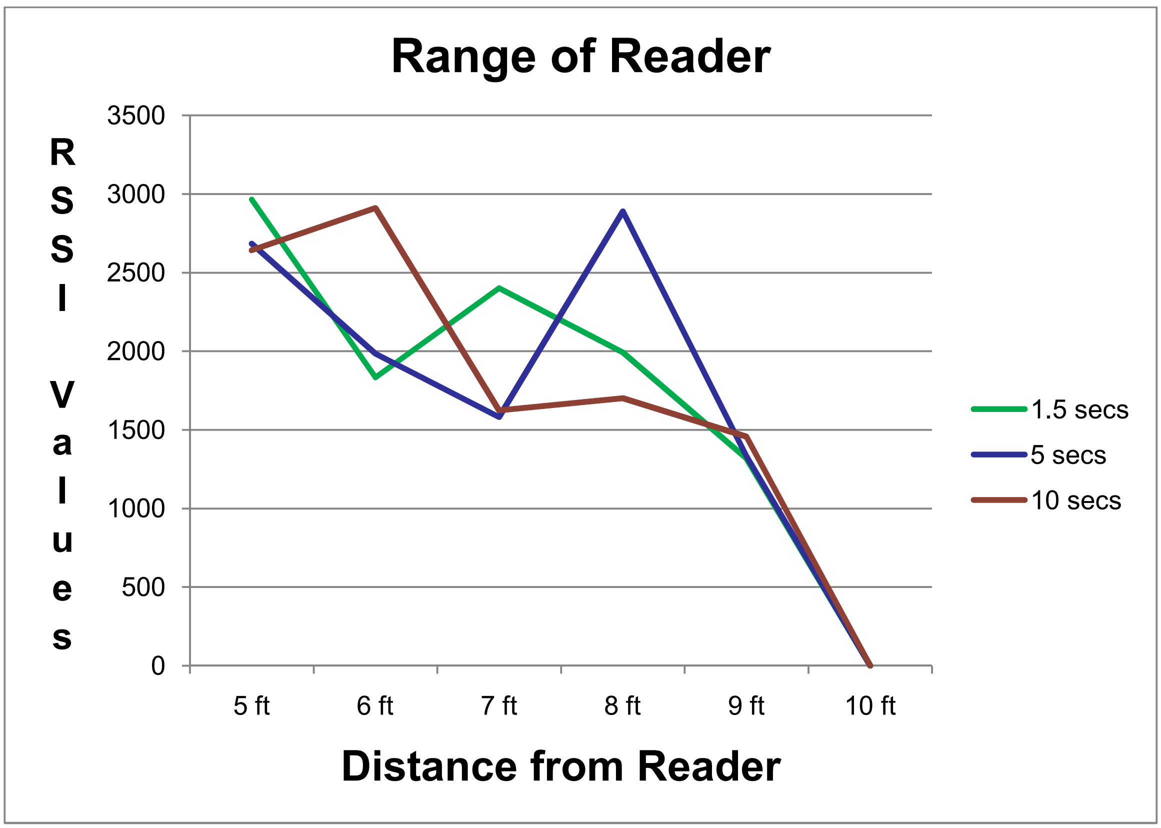

Figure 15: Effect of distance of tag from reader on RSSI values

As seen in fig.15 ideally the RSSI values should decrease (shown by red curve in the graph) as the distance of the tag from the reader increases (for different data collection time spans i.e. 1.5, 5 and 10 seconds) but that doesn't happen practically and the change in RSSI values is very random as the tag keeps moving away from the reader. This leads us to the conclusion that RSSI value of a detected tag by itself cannot be considered to be a direct measure of its distance from the reader and has to be taken as parameter for a particular localization algorithm. The above graph also tells us the range of the reader which is around 9 – 10 ft (around 3m).

5.2 Fluctuations in RSSI values when tag is stationary at a particular place

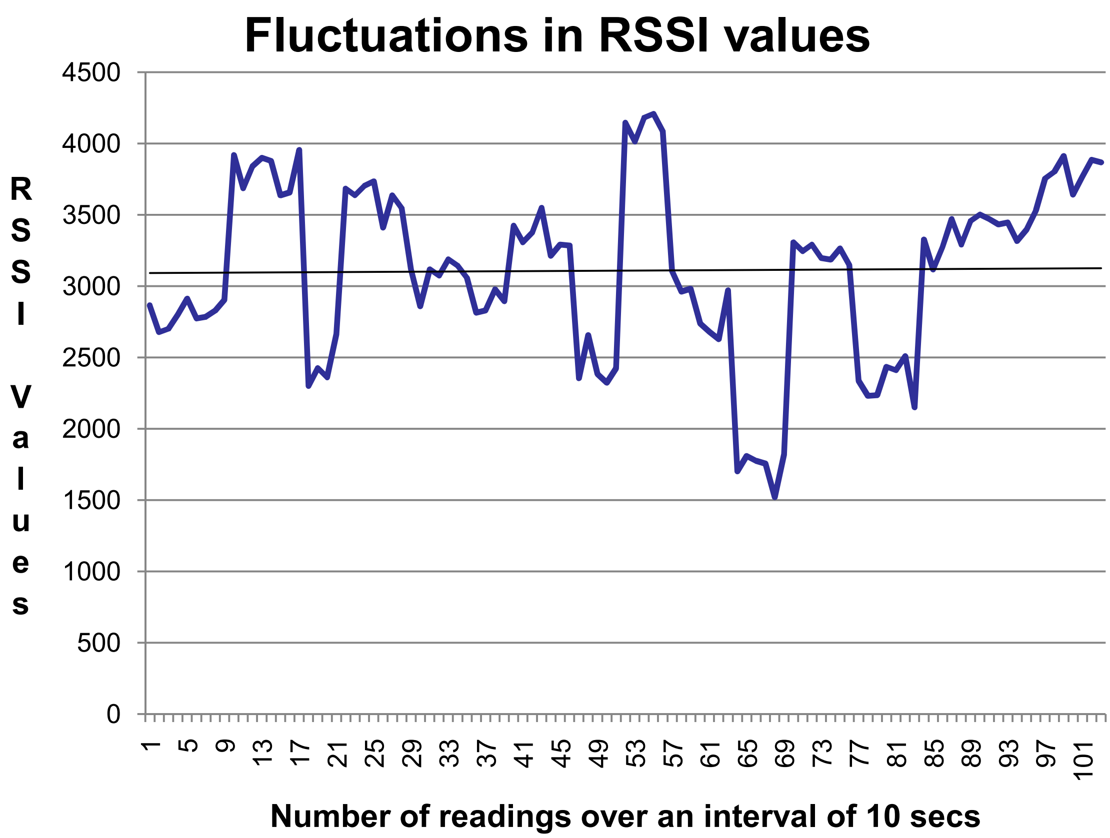

Figure 16: Fluctuations in RSSI values when tag is stationary at a particular place

From fig.16 we see that when the tag is kept stationary at a particular position from the reader, the registered RSSI values (shown by blue curve in graph) are fluctuating as compared to an ideal scenario (shown by the black line in graph) wherein the values would be more or less constant. Here we can conclude that we cannot use RSSI values registered on the fly as a localization parameter due to its fluctuations. A feasible solution for this problem would be collecting the RSSI values over a small period of time (based on accuracy of the system) and using a mean of all those values as a parameter. We have used this proposed solution in our algorithm.

5.3 Effect on RSSI values due to tags stuck on various materials

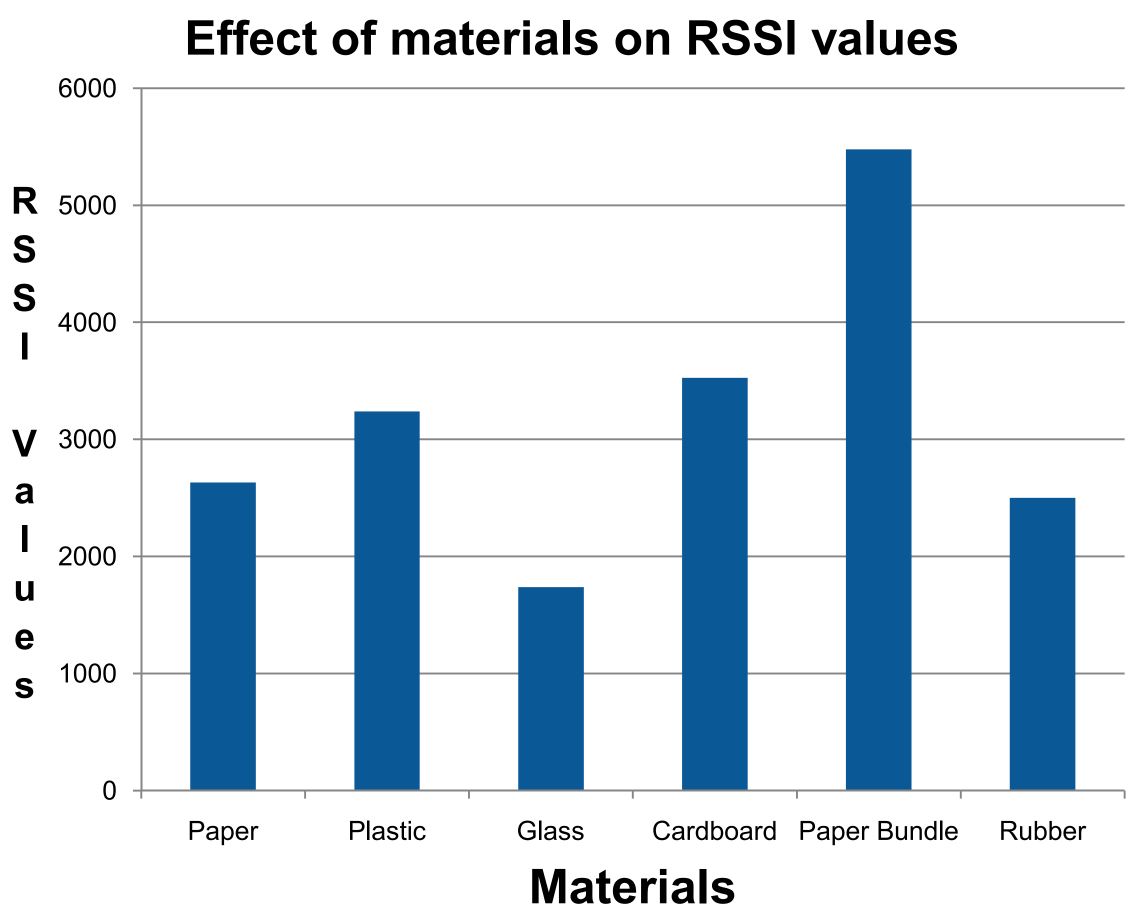

Figure 17: Effect on RSSI values due to tags stuck on various materials

Materials, on which tags are stuck, cause a significant variation in registered RSSI values as seen in fig.17.

5.4 RSSI Heat map

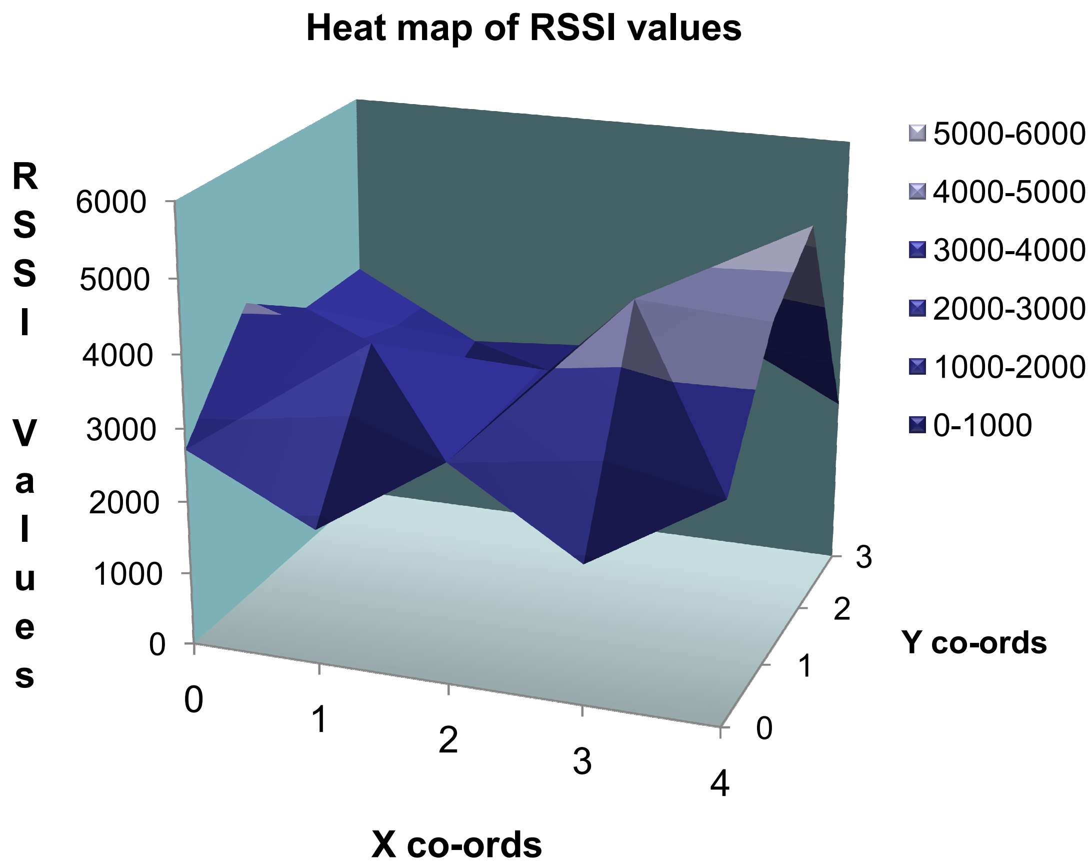

Figure 18: RSSI Heat map

The heat map in fig.18 indicates that the distribution of RSSI values over the entire field area is very random and is dependent on environmental factors as well. The result has been based on the collection of RSSI values by placing the tag at reference positions for 1.5 seconds (justified in next section).

5.5 Effect of multiple tags in a 15x15 in area on RSSI values

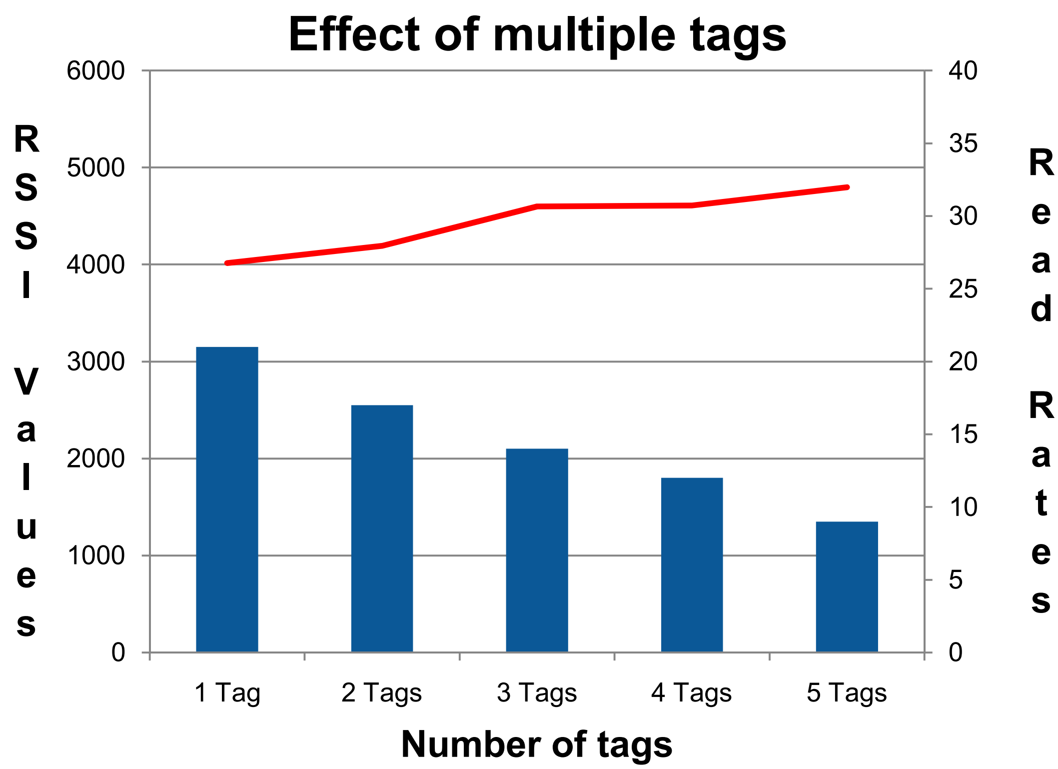

Figure 19: Effect of multiple tags in a 15x15 in area on RSSI values

As shown in fig.19, the RSSI values for a particular tag remain more or less constant when up to 5 tags are placed in a 15x15 inch block but a decrease in the read rates can be observed. This observation is of utmost importance in our application where in there can be a situation where multiple tags need to be detected simultaneously.

5.6 Mean Error for different data collection time spans.

As discussed earlier since the RSSI values fluctuate even though they are stationary at a particular position it is not advisable to take RSSI values on the fly and feed them to the algorithm. We decided to collect them over a specified period of time and feed their mean value as an input to the algorithm. With this purpose we tested the algorithm's mean error rate for different data collection time spans and found a time span of 1.5 seconds to be the one for which the error rates were acceptable as shown in figure 20 below. The choice of the optimum time span was based on finding the right balance between having a sufficiently large time span to get more RSSI values so that we could have a consistent average and having a time span as close to real time to cater to the speeds of object movement in our application.

Figure 20: Mean Error for different data collection time spans.

6. Results and Observations

6.1 Mean Localization Error when reference positions are 10 inches apart.

The first experiment was carried out by having the reference positions about 10 inches apart and by using just two antennas in orthogonal positions (fig.10).

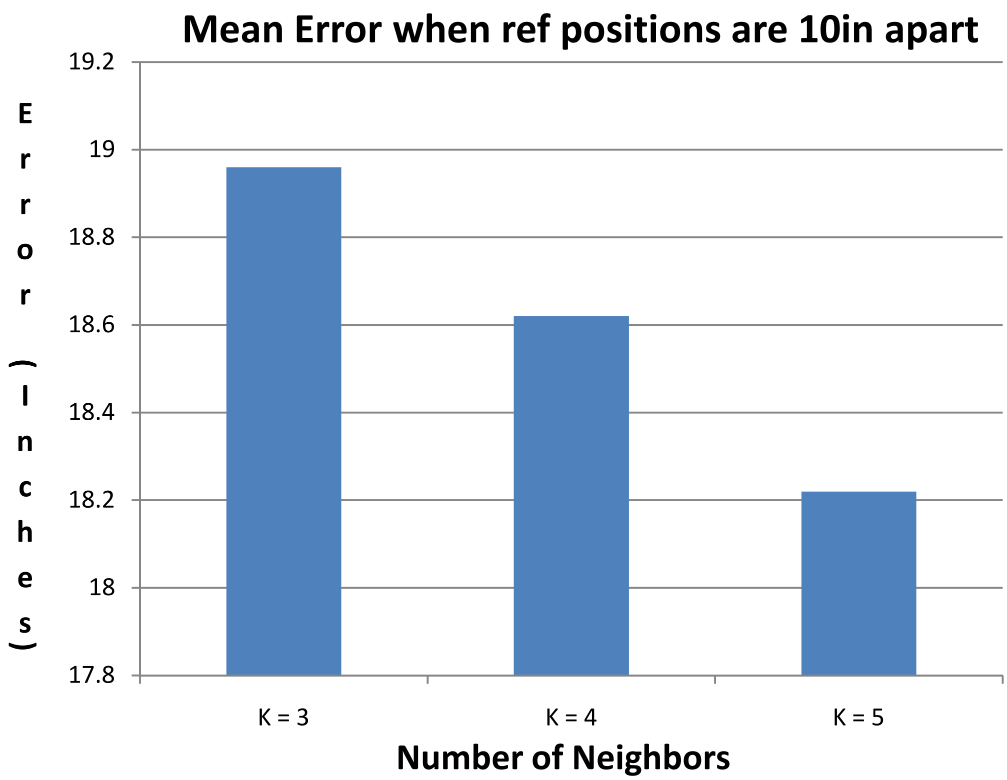

Figure 21: Mean Localization Error when reference positions are 10 inches apart

As seen above the mean error for this configuration was between 18 – 19 inches for 3, 4 or 5 neighbors. As evident this result was not at all satisfactory so we decided to increase the distance between reference positions to 15 inches as it has been seen that difference in RSSI values between two blocks 1m apart is much higher than values between blocks few cm apart [16].

6.2 Mean Localization Error when reference positions are 15 inches apart.

The next configuration as suggested above was having the reference position spaced 15 inches apart with the same set up of antennas i.e. they were orthogonally placed.

Figure 22: Mean Localization Error when reference positions are 15 inches apart

Fig.22 shows that increasing the distance between reference positions does improve the accuracy of localization. The error for this configuration was recorded to be around 15 inches which is about 3 inches better than the earlier configuration.

6.3 Mean Localization Error when reference positions are 15 inches apart and using 3 antennas.

The next step in improving the accuracy was to increase the number of antennas in the field area. We decided to use 3 antennas instead of 2 antennas and placed them as shown in fig.11.

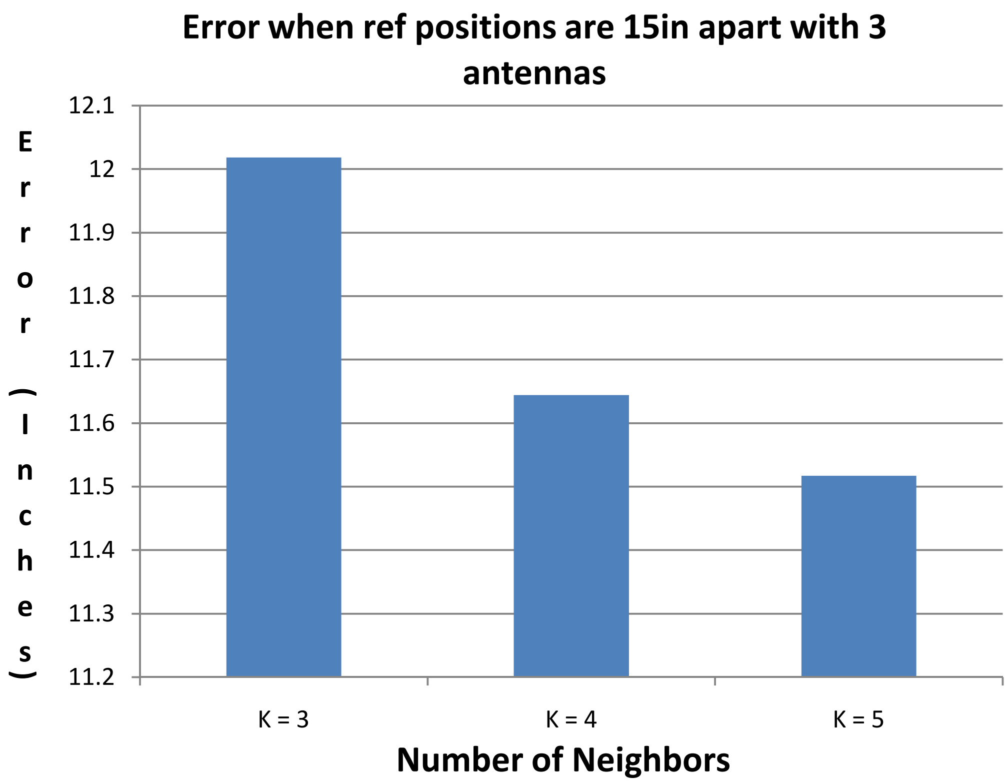

Figure 23: Mean Error when reference positions are 15 inches apart and using 3 antennas

This particular configuration recorded an error of around 11.5 – 12 inches with K=5 fetching the best result. This again was a good 3 inches better than the configuration having 2 antennas.

6.4 Localization Error when reference positions are 15 inches apart and using 3 antennas (2 symmetric and 1 vertical antenna).

While tracking surgical instruments in a hospital's emergency room, the antennas are going to be occluded by a number of humans walking around in the field area. There can be a situation where all the 3 wall antennas are blocked and no RSSI values are fetched by any of the antennas. In that case it would be impossible to track the object tag under the current set-up. So, we needed to place an antenna at a position where no matter what we would get some RSSI value at least so that we have an input parameter for our algorithm to make an approximate localization decision. Hence we decided to place 1 antenna on the ceiling and the remaining 2 on opposite walls as shown in fig 12.

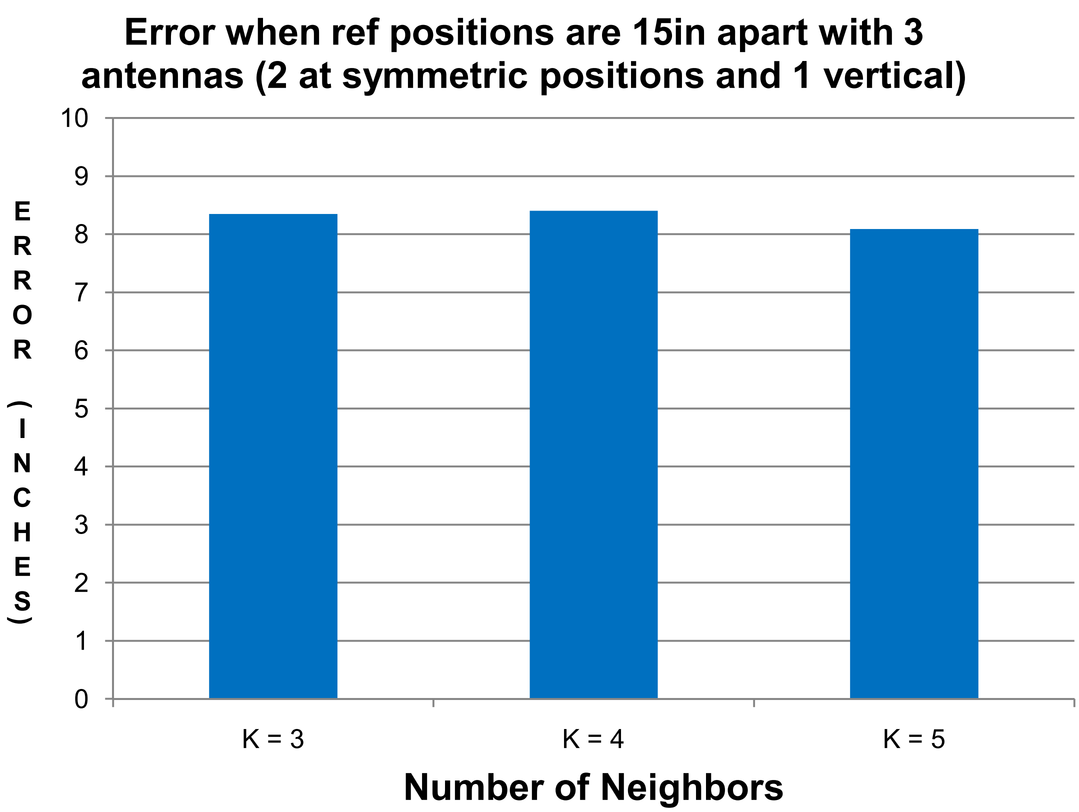

Figure 24: Mean Error when reference positions are 15 inches apart and using 3 antennas (2 symmetric and 1 vertical antenna)

The error for this configuration improved by around 4 inches and at the same time it proved to be a good enough solution to the problem of human occlusion as illustrated by the following results.

6.5 Localization Error when reference positions are 15 inches apart and using 3 antennas (2 symmetric and 1 vertical antenna) where one or two of the antennas are blocked.

As discussed above, testing the algorithm for human occlusion was of prime importance in our application. We started off with one of us standing in front of each of the wall antennas and blocking it. The results for the same are displayed below.

6.5.1 Error on blocking Antenna 0 in fig 12

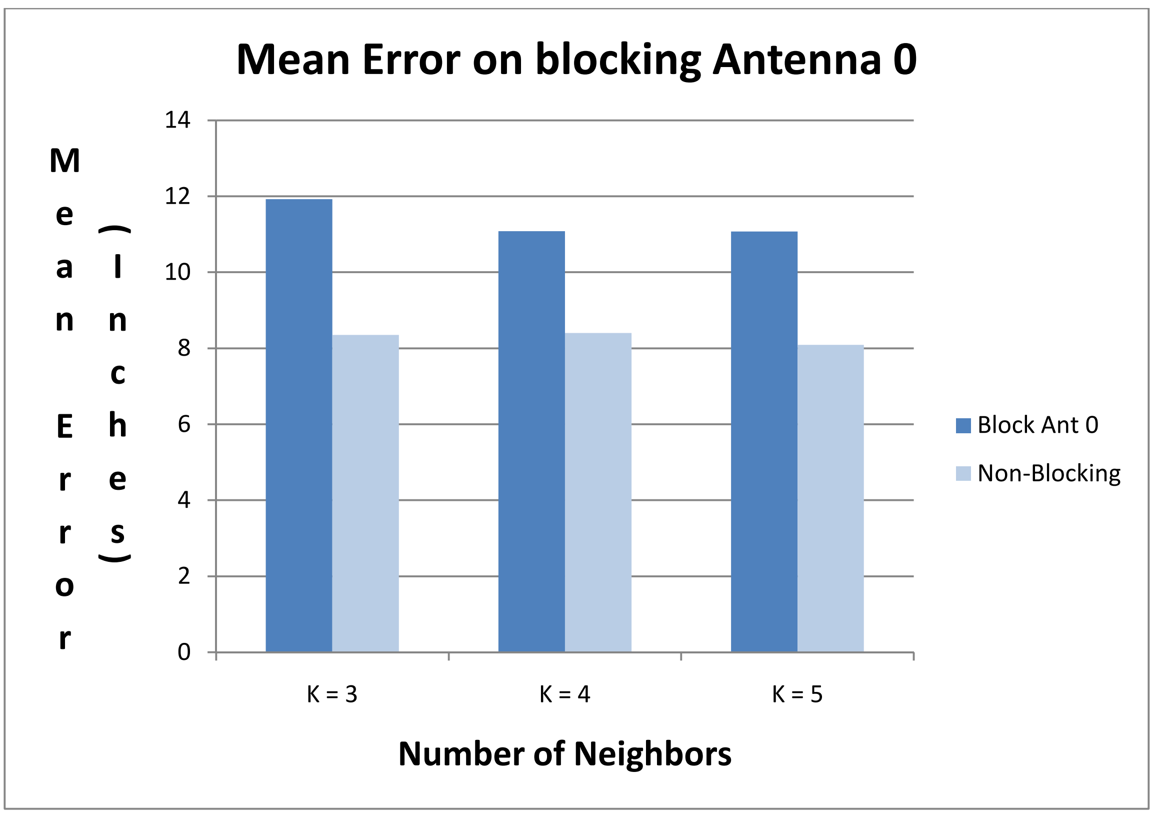

Figure 25: Error on blocking Antenna 0 in fig 12

6.5.2 Error on blocking Antenna 2 in fig 12

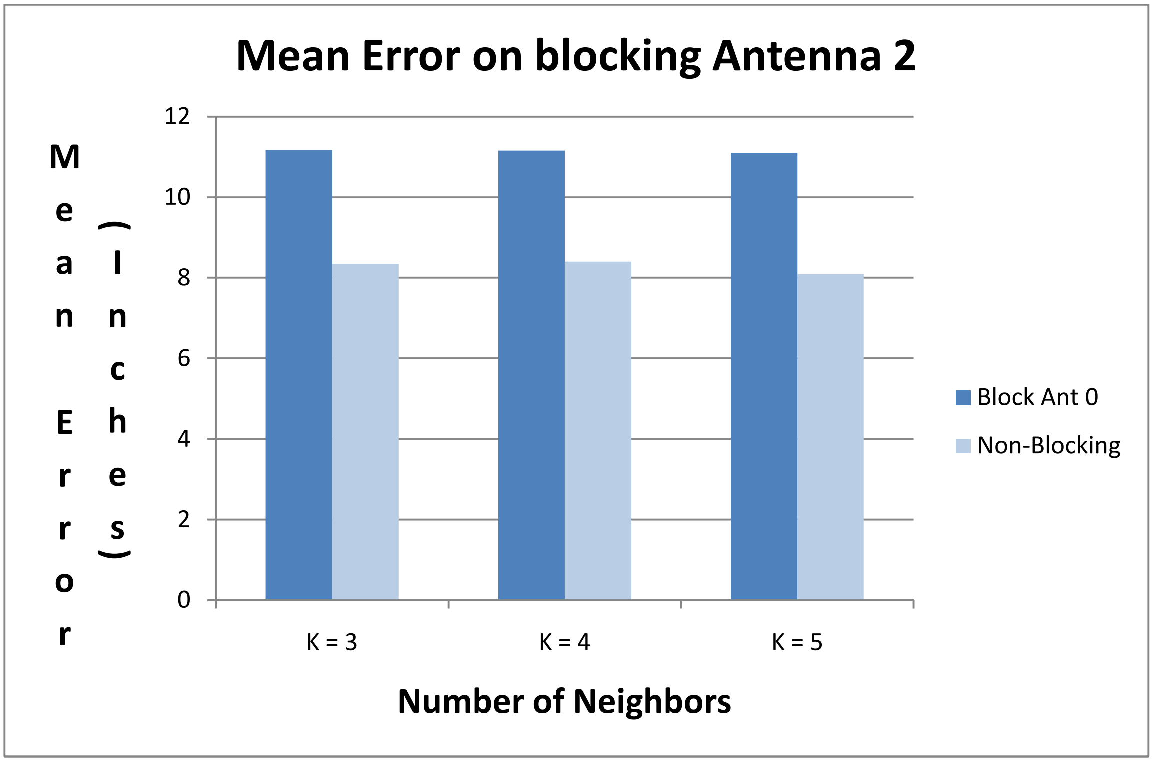

Figure 26: Error on blocking Antenna 2 in fig 12

6.5.3 Error on blocking Antenna 0 & 2 in fig 12

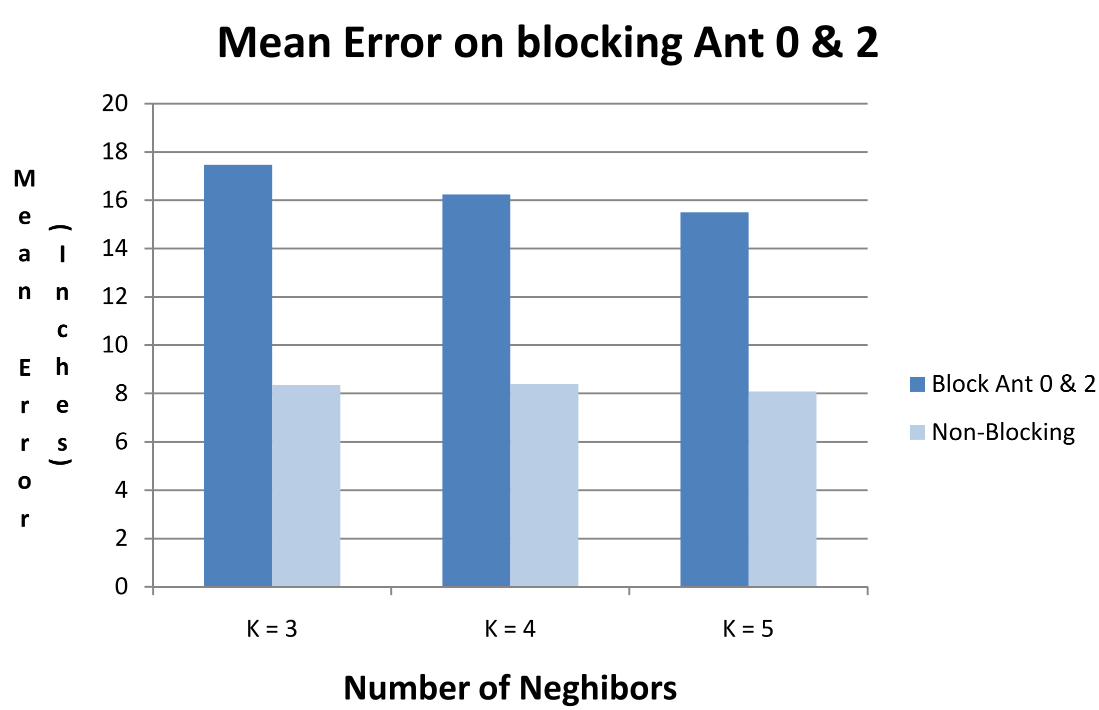

Figure 27: Error on blocking Antenna 0 & 2 in fig 12

From the above 3 figures (fig 25, fig 26 & fig 27) we see that when either antenna 0 or antenna 2 is blocked we get uninterrupted RSSI values from the other 2 antennas which decide the algorithm output. In both the cases the error was around 11 inches which is pretty good for a tag which itself is about 5 inches long. Whereas when both antennas 0 & 2 were blocked the error jumped to about 15.5 inches which was the best result for K = 5 as against 16 inches (K = 4) and 17.5 inches (K = 3). Although the error is a bit on the higher side but at least it provides us with some idea about the tag's position as compared to nothing in the earlier configuration with 3 wall antennas.

Apart from mean error another aspect that we thought would be worth considering is how consistently the algorithm predicts the co-ordinates to be within a certain specified error rate. This led us to generating a histogram of all the calculated values i.e. 60 experiments (10 experiments in each block)

6.5.4 Histogram of 60 experiments for 3, 4 and 5 neighbors

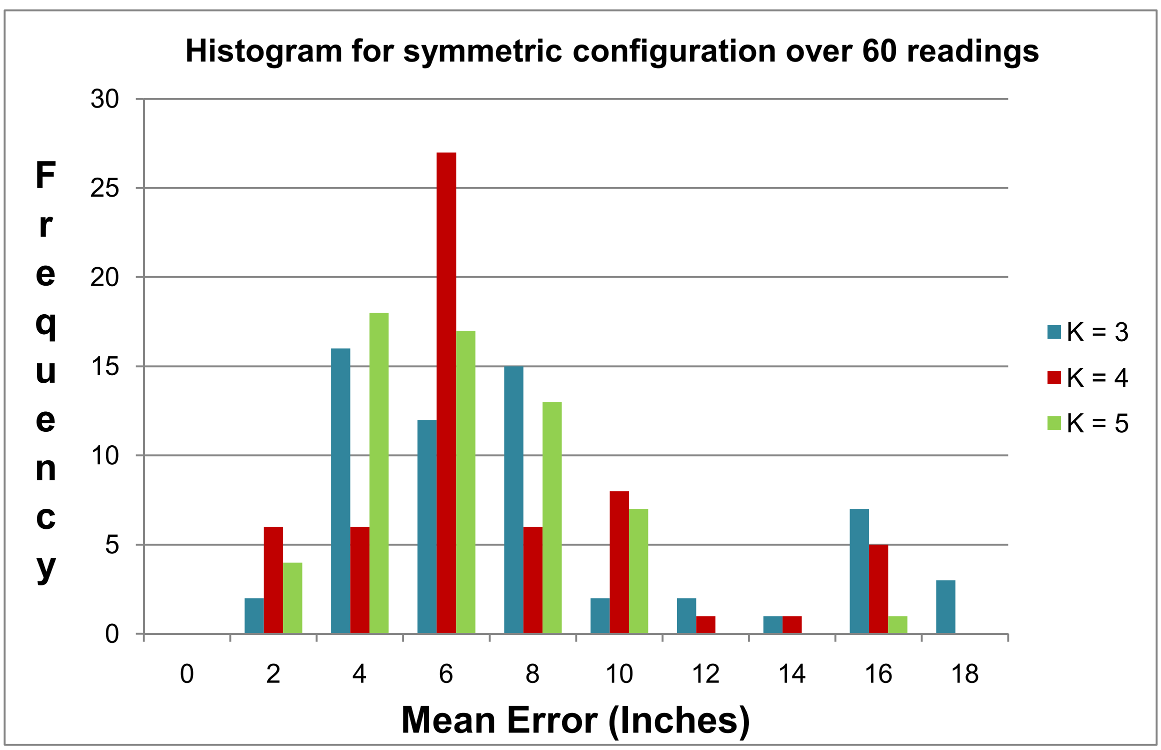

Figure 28: Histogram of 60 experiments for (K = 3, 4, & 5) – symmetric configuration

On analyzing the above histogram, if we set a maximum error threshold of 10 inches, the following observations can be made.

- 83.33% of experiments are in acceptable range (< 10 inches) for K = 3
- 91.67% of experiments are in acceptable range (< 10 inches) for K = 4
- 98.33% of experiments are in acceptable range (< 10 inches) for K = 5

Thus K = 5 proves to be the best configuration with an error of 8 inches and an accuracy of 98.33% when 3 antennas (2 symmetrically placed and 1 vertical antenna) are used.

6.6 Localization Error when reference positions are 15 inches apart and using 3 antennas (2 orthogonally placed and 1 vertical antenna).

After testing the algorithm with the above configuration we thought it would be interesting to test the algorithm with an orthogonal configuration where 3 antennas cover all the three dimensions. So we placed 2 antennas on the wall in orthogonal positions and the third antenna on the ceiling as shown in fig.13.

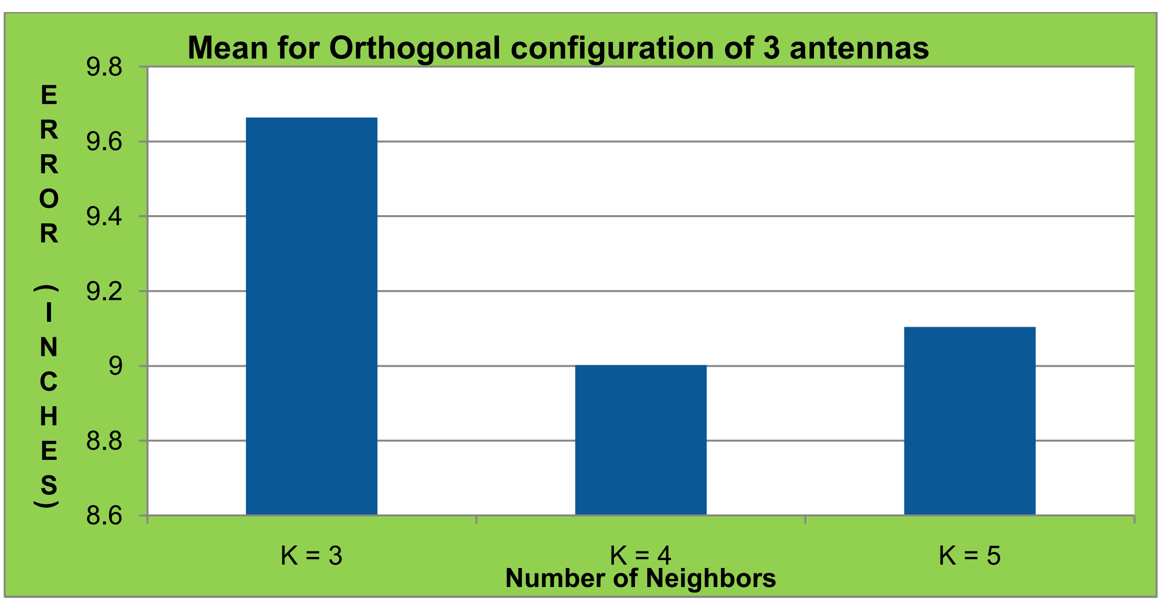

Figure 29: Localization Error when reference positions are 15 inches apart and using 3 antennas (2 orthogonally placed and 1 vertical antenna)

The error recorded in this configuration was an inch higher than the symmetric configuration but the error was found to be consistently within acceptable range (< 10in) making it comparatively more stable as shown in the histograms below.

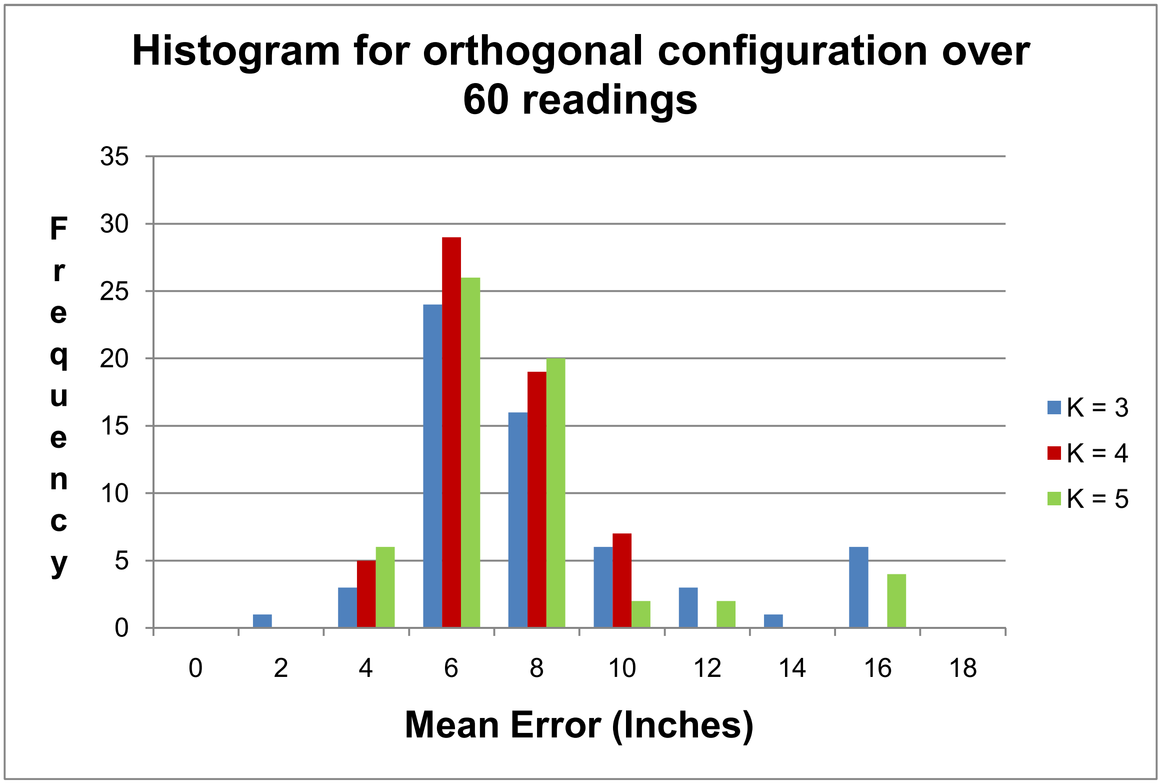

Figure 30: Histogram of 60 experiments for 5 neighbors (K = 3) – orthogonal configuration

The following observations can be listed from the above histograms

90% of experiments are in acceptable range (< 10 inches) for K = 3

100% of experiments are in acceptable range (< 10 inches) for K = 4

98.33% of experiments are in acceptable range (< 10 inches) for K = 5

Hence for this particular combination K = 4 would be a better choice with a mean error of 9 inches and an accuracy of 100%.

6.7 Mean error when 2 people are standing and then 3 people roaming in the field area for orthogonal configuration.

This is one of the most significant experiments for our application as explained earlier where humans will be standing or moving randomly as shown in figures below. The results for these experiments are discussed below as well.

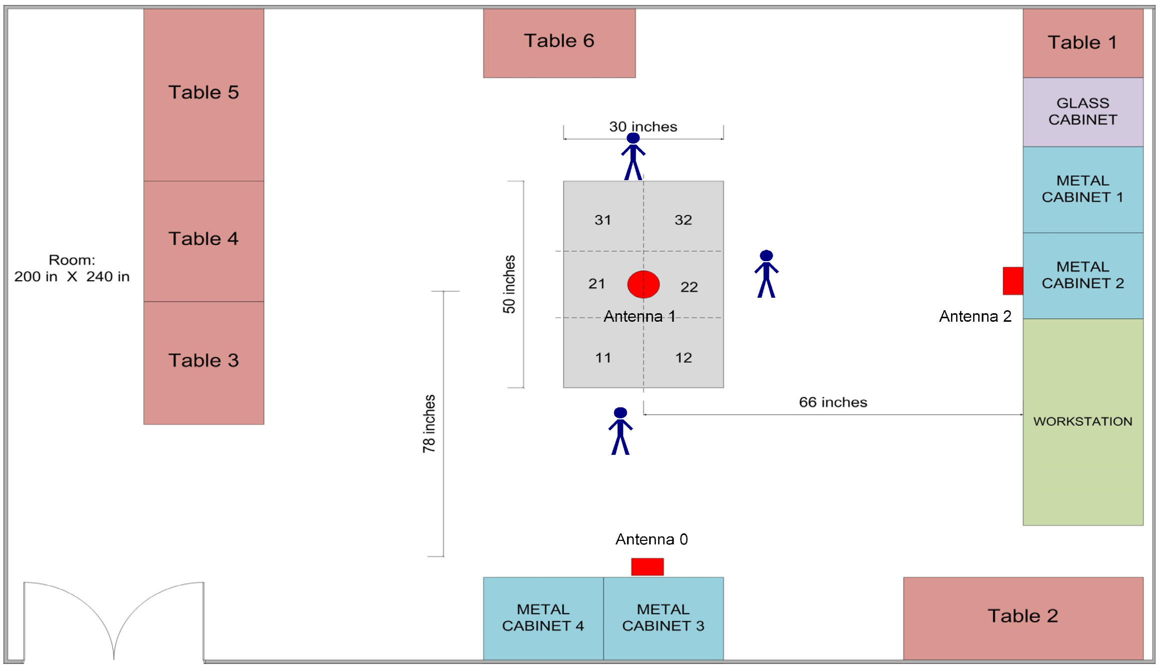

Figure 31: Floor Plan for set-up with orthogonal config. and 3 people standing randomly in the area

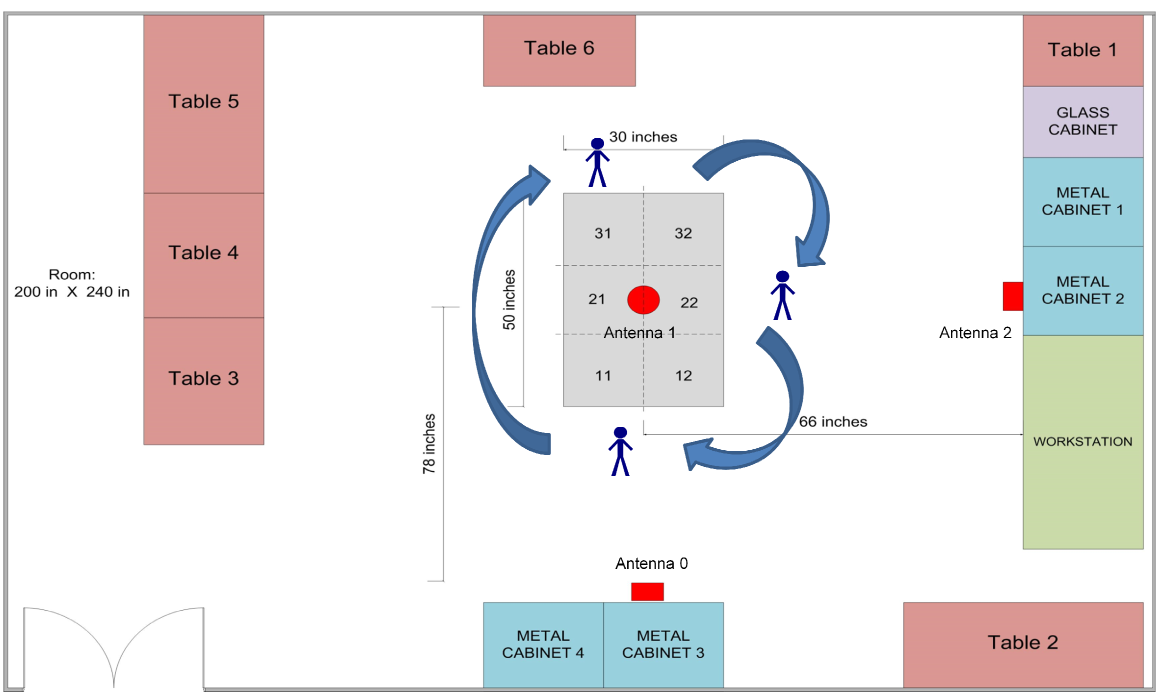

Figure 32: Floor Plan for set-up with orthogonal config. and 3 people roaming randomly in the area

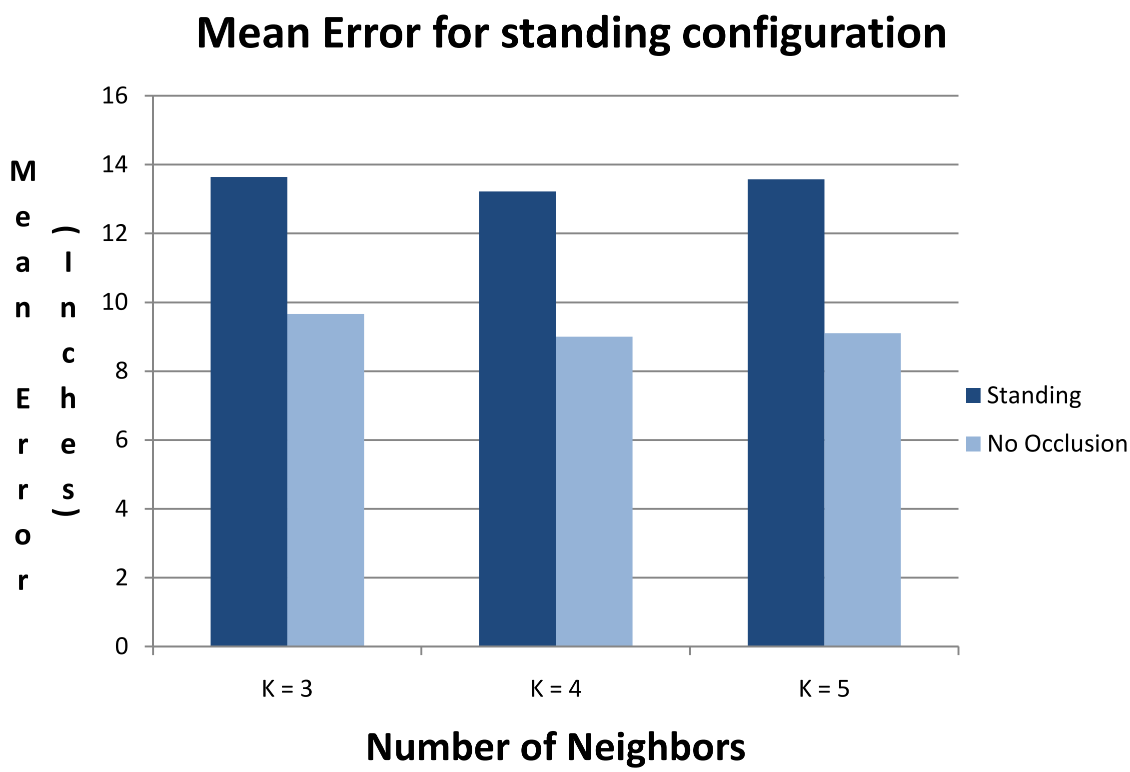

Figure 33: Mean Error for standing configuration

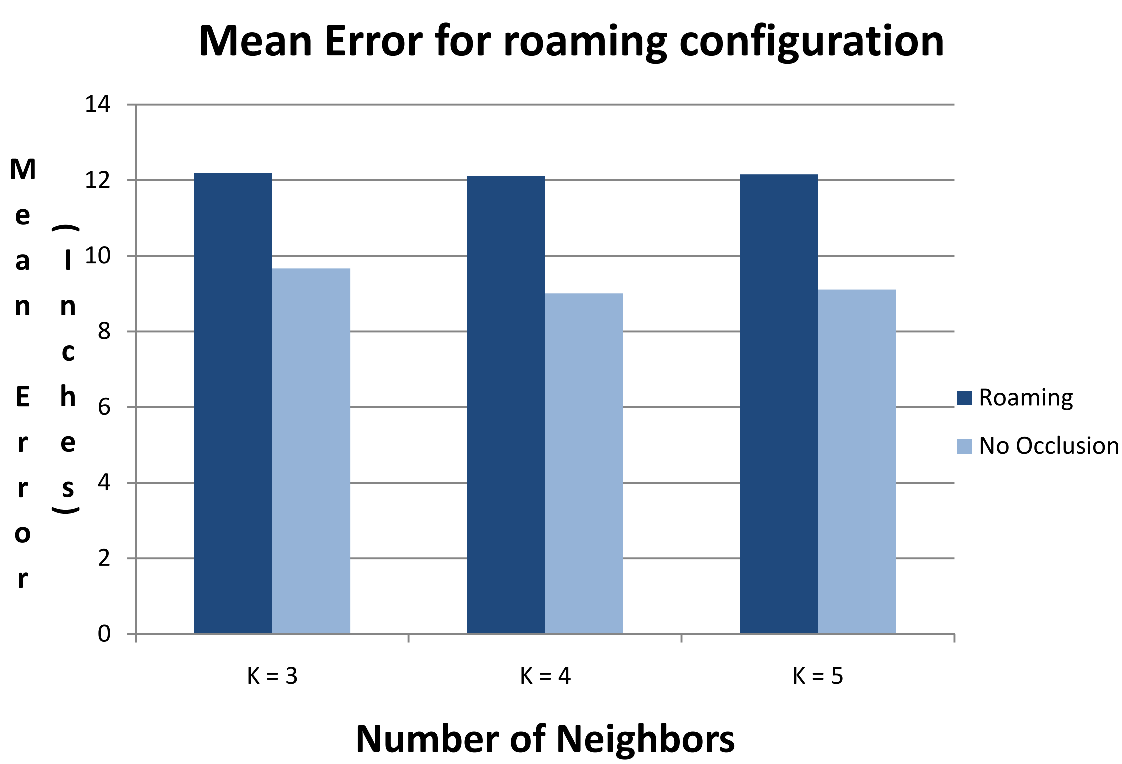

Figure 34: Mean Error for roaming configuration

Fig.33 and fig.34 reveal that the mean error when 2 people are standing in the field area is around 13 inches and the mean error when 3 people are randomly moving in the field area is around 12 inches. The reason for this can be explained as the fact that when 2 people are standing almost all RSSI values are interrupted whereas when people are randomly moving at least some of the values are uninterrupted giving better results.

6.8 Mean error of an object tag when up to 5 tags are placed within a 30 x 30 inch block.

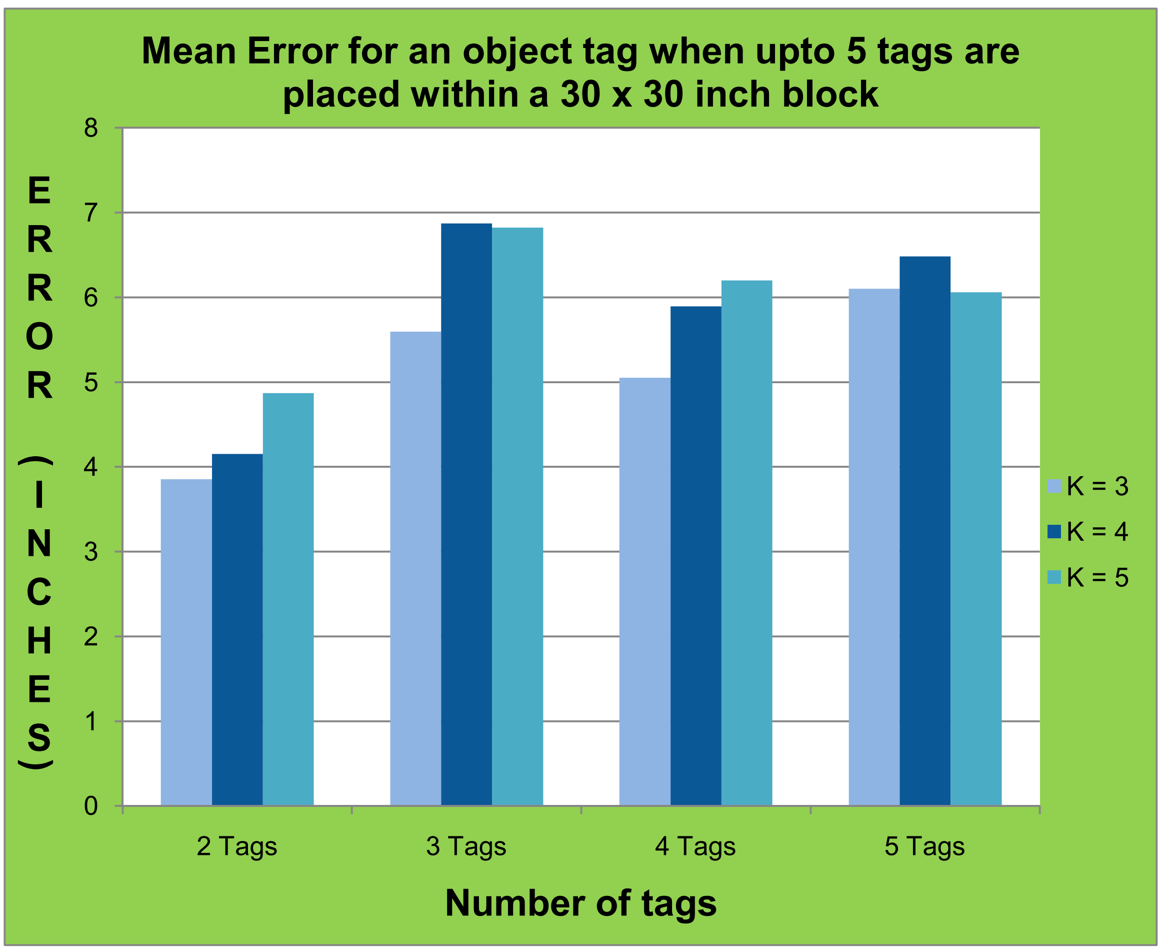

Figure 35: Mean Error for an object tag when upto 5 tags are placed within a 30 x 30 inch block

Shown above in fig.35 are the mean errors when 2 to 5 tags are placed in a 30 x 30 inch block. The mean error for all these tests falls within 7 inches making the algorithm feasible for detection of at least 5 tags which is an important aspect of our application where multiple tags might have to be detected simultaneously.

Figure 36: Mean error for all 5 tags placed within 30x30 inch block

Fig.36 shows the mean error for all 5 tags when placed in a 30 x 30 inch block over 10 experiments per tag.

6.9 Mean error for object localization over a 9 x 9 feet area under the orthogonal antenna configuration with center and edge test points.

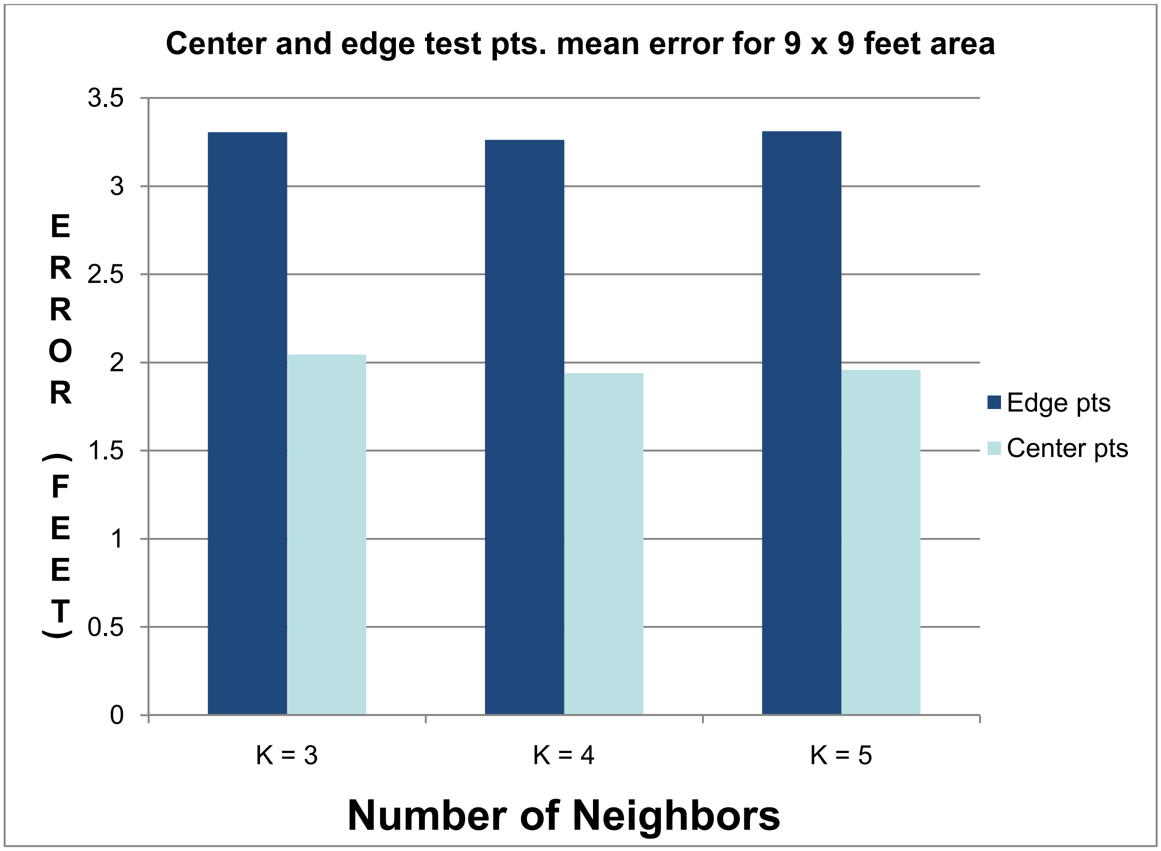

Figure 37: Mean error for object localization over a 9 x 9 feet area for center test points.

Shown above in figure, are the mean errors for the center and edge test points when the algorithm is applied over a 9 x 9 ft area under the orthogonal antenna configuration. The results show that the algorithm works just as fine when it applied over a larger area. A mean error of around 1.9 feet was recorded for the test points at the center of each 3 x 3 feet block whereas it increased to around 3.2 feet for the test points located near the edges of each block.

The reason for increase in the mean error for the edge test points can be tracked down to the fact that just 3 antennas are not sufficient to cover the entire area resulting in the tags not being detected by one or more antennas at edge test points. Also the number of RF reflected RF waves from the surroundings are found to be more at the edge points than at the center points affecting the recorded RSSI values.

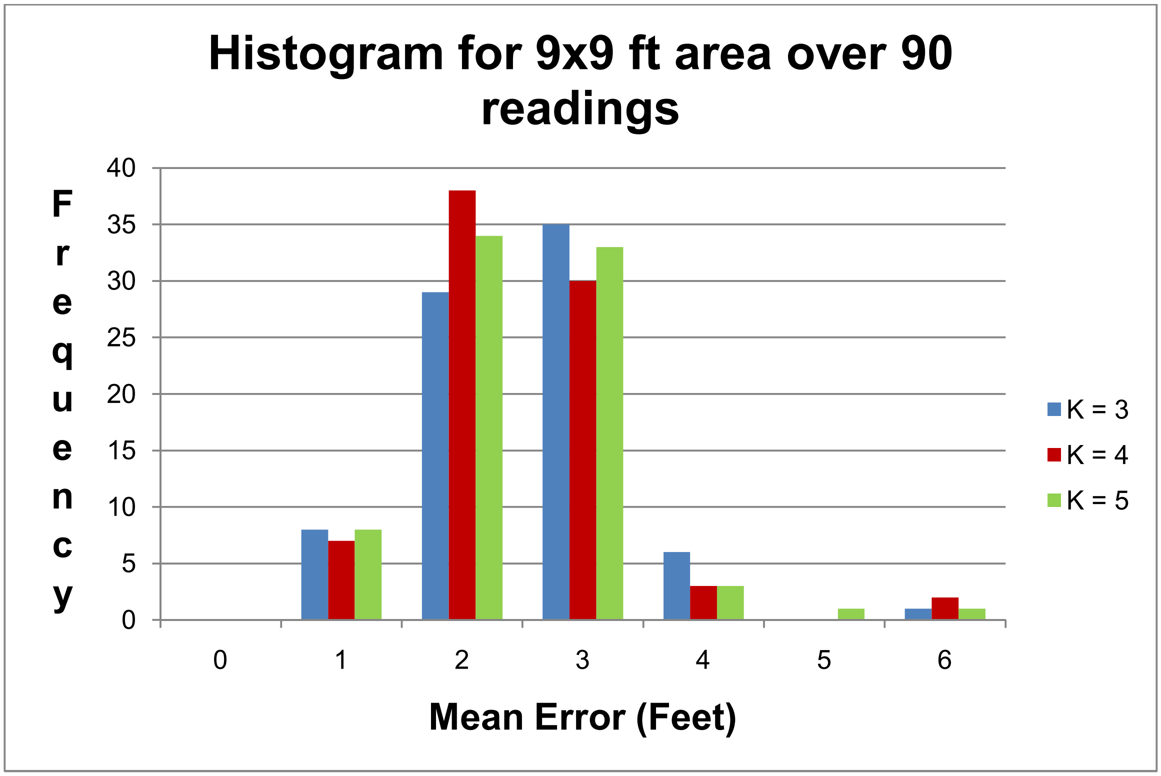

Figure 38: Histogram for 9x9 ft area over 90 readings

The following observations can be listed from the above histograms

92.2% of experiments are in acceptable range (< 3 feet) for K = 3.

95% of experiments are in acceptable range (< 3 feet) for K = 4.

93% of experiments are in acceptable range (< 3 feet) for K = 5.

6.10 Time Complexity

The time complexity of the algorithm depends on the number of reference positions (r), number of inputs i.e. number of antennas (n) and the number of neighbors (k).

For a fixed number of neighbors, if we consider 'r' reference positions and 'n' inputs, the algorithm will have a complexity given by $\Theta(n + r \log r)$, as the algorithm requires sorting of the reference points to find the ‘k’ nearest ones and it increases linearly as the number of inputs is increased.

Also there is a linear increase in the complexity of the algorithm with an increase in the number of neighbors being considered. Hence, k=3 will have a better complexity as compared to k=4 and so on.

7. Discussion of Results

In this section we discuss how our work tries to answer research questions raised at the beginning of this thesis.

1. Scalability:

   The system works fine for a large area of 9x9 feet by providing a mean error which is less than 15% of the maximum possible error.

2. Accuracy:

   As mentioned above error rates comparable to 15% of the maximum possible error are achieved under ideal conditions whereas it goes up to about 18% under human occlusion which is still better that are pre-defined tolerance of 20%.

3. Cost:

   The system costs in the order of a few thousands which is within our specified budget limits.

4. Performance:

   The algorithm takes an average of about 30ms for the system specified below:

   System:

   Microsoft Windows XP Professional x64 Edition

   Version 2003 Service Pack 2

   Computer:

   Intel(R) Xeon(R) CPU

   E5420 @2.50GHz

   2.49GHz, 4.00GB of RAM

To specify the approximate time frame of real time, let us assume that the tag moves by about 5 inches in an infinitesimally short time. And now if the person carrying the equipment is moving at speeds comparable to fast walking speeds i.e. 4.5mph then the time elapsed can be calculated and is found to be 63ms. Thus we can see that the results are comparable.

5. Multiple tag detection:

   The algorithm detects multiple tags with mean error of around 9 inches as compared to the detection of a single tag that gives a mean error of 8 inches.

6. Overcoming Human Occlusion:

   A maximum error rate of 18% is achieved which is again lower than the pre-defined error rate of 20% of the maximum possible error.

The above discussion reveals that we have provided satisfactory solutions to the research questions we faced while designing the system.

Variance and Standard Deviation

We performed 60 experiments for each configuration and have considered the mean of the error rates for all those 60 readings as a metric to assess algorithm accuracy. Since we are reporting the mean errors itself, the calculation of variance and standard deviation is no longer logical fundamentally. However, the range of error for all the 60 readings can be specified with the algorithm yielding a minimum mean error of 2 inches and a maximum of 9 inches for the K = 4 orthogonal configuration. The distribution for all the 60 readings has been plotted earlier in the histograms in figs. 28, 30 and 38.

8. Conclusion

Config	Mean Error			Consistency (%)			Mean Error as percentage of maximum error (%)		
	K = 3	K = 4	K = 5	K = 3	K = 4	K = 5	K = 3	K = 4	K = 5
Symmetric	8.23 in	8.32 in	8.19 in	83.33	91.67	98.33	13.72	13.87	13.65
Orthogonal	9.65 in	9 in	9.154 in	90	100	98.33	16.08	15	15.26
9x9 feet	2.045 ft	1.94 ft	1.957 ft	92.2	95	93.1	15.73	14.91	15.05

The table above clearly indicates that the Weighted K-Nearest Neighbor algorithm works well for our application under the mentioned constraints. The highlighted fields indicate the best results for that particular configuration. Of all the configurations discussed in chapter 5, the two most comparable ones are K = 5 for symmetrical configuration and K = 4 for orthogonal configuration. But the latter scores as it is more stable and has a better time complexity. The only drawback would be a mean error of 9 inches as against 8 inches for the former but again a trade-off of 1 inch doesn't really matter when all other factors are considered.

This is further confirmed by the results obtained for the center test points when the algorithm is applied to a larger area. However, the edge test points show a significant decrease in accuracy but this is primarily due to bad area coverage apart from some other factors like reflection, refraction, scattering, interference, etc. This is the best we could achieve in terms of accuracy based on the resources available to us. Better results can be obtained by employing more readers and antennas to provide better area coverage.

Also, the algorithm efficiently tracks objects even when you have multiple tags in the field or humans wandering randomly.

Shown below is a table describing how the WKNN algorithm compares with the other existing localization algorithms.

Localization Technique	Mean Error as a percentage of Maximum Possible Error
Kalman Filtering	50%
Bayes Decision Approach	20%
LANDMARC	17%
Maximum Likelihood Classifier (Gaussian)	16%
Machine Learning Reference	13%
WKNN	15%

As seen from the table above, the WKNN algorithm fares better than most other algorithms except the Machine Learning Reference algorithm (MLR) which itself is an offshoot of WKNN since it merges both the Bayesian Decision and the WKNN approach. Thereby there are concerns about its deployment as it requires extensive data collection, maintenance of significant result history and also it may not be close enough to real time.

9. Future Work

While testing our algorithm we found that the received signal strength (RSSI) values vary depending on the material on which the tag is attached. Finding an alternative or improving our algorithm to accommodate detection of different materials would provide for interesting research. Better results can be achieved when the algorithm is implemented over a larger area if more readers and antennas are employed to improve area coverage as we worked within our resources. Also the RSSI values depend a lot on the orientation of tags in the field area; a solution to this problem is a must as there can be no restriction on how a particular instrument is supposed to be held while live-tracking. Tracking of moving tags can also be the next step in improving the system as we have analyzed just the localization of static tags in our work. Finally, we think read rates can be employed along with taking the mean of RSSI values over a period of time to reduce the data collection time span and make the system as close to real-time as possible.

10. References

1) www.rfidjournal.com Last visited: 26 July 2010

2) Patrik Fuhrer, Dominique Guinard - University of Fribourg, “Building a Smart Hospital using RFID technologies”, Proceedings of 1st European Conference on eHealth (ECEH06), Fribourg, Switzerland, October 12 - 13, 2006.

3) Patrick J. Sweeney II, “RFID for dummies”, Wiley Publishing Inc,2005.

4) Klaus Finkenzeller, “RFID handbook: fundamentals and applications in contactless smart cards and identification”, John Wiley & Sons Inc,2nd edition,2003.

5) Mathieu Bouet, Aldri L. dos Santos, “RFID Tags: Positioning Principles and Localization Techniques”, Proceedings of Wireless days,2008. WD ’08. 1st IFIP.

6) Junyi Zhou, Jing Shi, “RFID localization algorithms and applications—a review”, Journal of Intelligent Manufacturing, Volume 20, Number 6, pps 695-707, July 2008.

7) Bulusu N., Heidemann J., Estrin D., “GPS-less Low-Cost Outdoor Localization for Very Small Devices” IEEE Personal Communications, Volume 7, Number 5, pps. 28-34, October 2000.

8) X.Li, K.Pahlavan , “Super-resolution TOA estimation with diversity for indoor geolocation”, IEEE Transactions On Wireless Communications, Volume 3, Number 1, pps 224 – 234, January 2004

9) K.Pahlavan, Li. Xinrong and J. P. Makela, “Indoor geolocation science and technology,” IEEE Communications Magazine, volume 40, issue 2, pp 112-11, Feb.2002.

10) D.Niculescu, B.Nath , “Ad Hoc Posiitoning System(APS) Using AOA”, Proceedings of IEEE Infocom Conference 2003.

11) Lionel M. Ni, Yunhao Liu, Yiu Cho Lau, Abhishek Patil, “LANDMARC: Indoor Location Sensing Using Active RFID”, Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003.

12) P. Bahl, A. Balachandran, and V. Padmanabhan, “Enhancements to the RADAR user location and tracking system”, Technical report, Microsoft Research, February 2000.

13) P. Prasithsangaree, P. Krishnamurthy, and P. K. Chrysanthis, “On indoor position location with wireless LANs,” Proceedings of 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC ’02), volume 2, pps.720–724, Lisboa, Portugal, September 2002.

14) T. Roos, P. Myllyma¨ki, H. Tirri, P. Misikangas and J. Sieva¨nen. “A Probabilistic Approach to WLAN User Location Estimation”. International Journal of Wireless Information Networks, Volume 9, Number 3, pps 155-164, July 2002

15) Bhatia, A., Mehta, B., & Gupta, R.: 2007, “Different Localization Techniques for Real Time Location Sensing using Passive RFID”

16) Ting Zhang, Zhenyong Chen, Yuanxin Ouyang, Jiuyue Hao, and Zhang Xiong. An improved rfid-based locating algorithm by eliminating diversity of active tags for indoor environment. The Computer Journal, Volume 52, Number 8, pps 902–909, 2009.

17) Steven Shepard , “RFID: radio frequency identification”, The McGraw Hill Companies Inc, 2005.

18) Sebastian van Hese, Master Thesis, University of Delft, The Netherlands – “Real-time localization and tracking of multiple people in a closed environment with facial detection using a multi-camera set up”, May 2008.

19) Alien 9900 RFID Reader Overview, http://www.alientechnology.com/docs/products/DS_ALR_9900.pdf

20) E. Inc. EPCTM generation 1 tag data standards, version 1.1 Revision 1.27, Technical report, EPCglobal Inc., 2005.

21) Matlab from Mathworks, http://www.mathworks.com/matlabcentral/ Last Visited – 10 August 2010

22) Alien Technology, EPCglobal Class 1 Gen 2 RFID Specification, White paper, 2005. http://whitepapers.techrepublic.com.com/abstract.aspx?docid=269751

23) Active RFID Information, Gao RFID, Last Visited – 12 July 2010 ‘http://www.gaorfid.com/index.php?main_page=index&cPath=130