

**Military Technical College  
Kobry El-Kobbah,  
Cairo, Egypt**



**11<sup>th</sup> International Conference  
on Electrical Engineering  
ICEENG 2018**

## **A Comparison Study of Indoor Localization Methods Using Available WI-FI Signals**

Abd Elgwad M. El Ashry\*, Ezz Eldin Farouk\*\* and Bassem I. Sheta\*\*

### **ABSTRACT**

Object localization in indoor areas strongly suffers from limitation of using GNSS (Global navigation satellite system-) systems due to low satellite availability and high signal attenuation. Nowadays, mobile devices such as personal computers and smart phones are emerging as a major key in today's computing platforms for indoor object localization systems due to the rapid developments in wireless communications and mobile computing. During last decade, many researchers have developed indoor localization systems through mobile devices using Wireless Fidelity (Wi-Fi) network signals with promising results and acceptable performance. In these Wi-Fi based localization systems, indoor positioning relies on different types of measurements including Time-Of-Arrival (TOA), Time-Difference-Of-Arrival (TDOA), Angle-Of-Arrival (AOA), and Received Signal Strength (RSS) of Wi-Fi signal.

In this paper, the algorithms and techniques used for the RSS-based localization systems using fingerprinting method which depends on matching the recorded offline RSS from nearby access points (AP) to the online RSS received by the user on the move is reviewed. A comparison of location fingerprinting methods involving deterministic method (k-nearest neighbor method and weighted k-nearest neighbor method), probabilistic methods by estimation of likelihood functions with several approaches (non-parametric and parametric) are also explained. The performance parameters of this comparative study include the two-dimensional root mean square error (2D-rms) which measures the localization accuracy. Moreover, the effect of increasing/decreasing the number of APs on the system accuracy is also discussed. The aim of this comparison is to announce which method can provide better performance than the others and under what conditions.

### **KEY WORDS:**

Wi-Fi signal, Indoor navigation, location fingerprinting, Received Signal Strength (RSS).

---

\* Egyptian Armed Forces.

## 1. Introduction

there are several limitations to the GNSS localization such as poor coverage of satellite signal especially in indoor environments. Consequently, several alternative approaches are introduced to provide efficient indoor localization where GNSS signal isn't available. These approaches include Wireless Local Area Net(WLAN, also known as WiFi), Bluetooth Low Energy (BLE), Radio Frequency Identification(RFID), ZigBee, Ultra Wideband Beacons (UWB), etc. Indoor localization has a wide spectrum of applications in both industrial and academic fields such as Real-Time Location Systems (RTLS), indoor mapping, indoor tracking, and other indoor localization applications. These companies include Apple, Google, Qualcomm, Intel, and Cisco along with start-ups that are paying attention to innovation in this field. These companies provide different localization accuracy depending on the technology they use. Another factor controlling the method used in indoor localization is the range provided by the method to gain localization accurately. For example, Bluetooth, UWB and RFID are used for short distance localization where ZigBee and WI-FI are used for middle distance localization. One of the major aspects that should be taken into consideration when investigating indoor localization is the practical aspects of the method used; such as system deployment, cost and time for system training. That is why Wi-Fi is one of the most effective approaches used for indoor localization since most buildings like super malls or office building have already deployed Wi-Fi hotspots that provide whole building coverage as network access point. And most commercial products, like phones, laptops and tablets, support Wi-Fi. That means the infrastructure cost and user device cost can be very low. With all the advantages, Wi-Fi is a mainstream technology in literatures for indoor localization.

Trilateration and fingerprinting are considered the most commonly used approaches for indoor localization. Trilateration is based on determining the location of the user using the geometry of triangles, while fingerprinting approach is based on comparing the RSS values to the ones stored in the database (radio map) to find the closest match. Fingerprinting has some advantages over trilateration method; it provides localization without prior knowledge of the access point location or the signal propagation model. Wi-Fi RSS is considered one of the most commonly approaches used for indoor localization.

Fingerprinting involves two phases, the off-line or the calibration phase and the on-line or the location estimation phase. During the offline phase, the system tabulates the signal strength received from the access points at selected locations in the area of interest, (RSS, Location), resulting in a so-called radio map. At online phase the system uses the signal strength samples received from the access points to "search" the radio map to estimate the user location.

In this paper, fingerprinting localization methods to estimate the user's location based on the online RSS measurements and the fingerprint database (Radio map) are investigated. These methods include deterministic and probabilistic approaches. The deterministic approach uses the average of the RSS time samples from each access point to estimate the location, whereas the probabilistic approach incorporates all the RSS time samples for the computation.

The rest of the paper is organized as follow: section-2, the preparation of RSS radio map is presented, section-3, the location estimation phase is covered from the deterministic method, section-4, the location estimation phase is covered from the probabilistic methods, section-5 summarizes the results

## 2. Preparation of RSS radio map

The radio map covers the area of interest and it holds the RSS values of the radio signal, which are collected as a function of the location. Signal to noise ratio (SNR) is also available, but it is often omitted because the RSS has a stronger correlation to the location than the SNR [3]. Design of the radio map is an important part of the location fingerprinting process and it constructs in offline phase, which contains position and signal feature of the position. The radio map can be modified or preprocessed before applying it during the location estimation phase. The motive for that can be the reduction of the required memory to store the radio map or the reduction of the computational cost of the location estimation. In addition, different location estimation methods apply different approximations from the fingerprints. Thus different ways to modify the raw data for the location estimation phase are discussed next.

During the calibration phase RSS values are measured at fixed locations for a certain period of time and stored in the radio map. The  $i_{th}$  element in the radio map has the form

$$M_i = (\text{position of CP}, \{ |a_{ij}| \mid j \in N_i \}, ) , \quad i = 1, \dots, M \quad (1)$$

where  $a_{ij}$  is the list of RSS values measured from access point  $AP_j$ . The set of Apsin range at the  $i_{th}$  calibration point is the set  $N_i$ , The number of samples measured from  $AP_j$  is the length of the list  $a_{ij}$ , denoted as  $|a_{ij}|$  and CP is a calibration point

The most common choice as a preprocessing method is to store only the mean of  $a_{ij}$  in the radio map. The  $i_{th}$  element in the radio map has the form

$$M_i = (\text{position of CP}, \{ |\bar{a}_{ij}| \mid j \in N_i \}, ) , \quad i = 1, \dots, M \quad (2)$$

where the mean of  $a_{ij}$  is:

$$\bar{a}_{ij} = \frac{1}{|a_{ij}|} \sum_1^{|a_{ij}|} a_{ij}^t$$

where  $|a_{ij}| = \text{length}(a_{ij})$  and  $a_{ij}^t$  is the  $t_{th}$  element of the list  $a_{ij}$ . The radio map of this form is used in Deterministic Approach.

The median value is the middle number in sorted order and is the mean of the middle two numbers in sorted order.

$$M_i = (\text{position of CP}, \{ |a| \mid j \in N_i \}, ) , \quad i = 1, \dots, M \quad (3)$$

where the value of  $(a)$  is the median value

The mode value is most recurrent value of the measured data at each calibration point.

$$M_i = (\text{position of CP}, \{ |m| \mid j \in N_i \}, ) , \quad i = 1, \dots, M \quad (4)$$

where the value of  $(m)$  is the mode value.

Storing just the means of the fingerprint does not give any information about the variation of the data. Fingerprints can be extended to store also the variance of the RSS samples [4]. Thus the  $i_{th}$  element in the radio map has the form

$$M_i = (\text{position of CP}, \{\bar{a}_{ij}, \widehat{\sigma}_{ij}^2 \mid j \in N_i\}, ), \quad i = 1, \dots, M \quad (5)$$

where the variance of  $\bar{a}_{ij}$  is

$$\widehat{\sigma}_{ij}^2 = \frac{1}{|a_{ij}|-1} \sum_1^{|a_{ij}|} (a_{ij}^t - \bar{a}_{ij})^2$$

and the sample size is assumed to be  $|a_{ij}| > 1$ . when the raw data is approximated with the Gaussian distribution.

The lists  $a_{ij}$  of RSS values can be presented as RSS histograms [5], thus the  $i_{th}$  element in the radio map has the form

$$M_i = (\text{position of CP}, \{H_{a_{ij}} \mid j \in N_i\}, ), \quad i = 1, \dots, M, \quad (6)$$

where the histograms are defined as the set of pairs,

$$H_{a_{ij}} = \{(b_t, h_t)\}$$

where  $b_t$  is the histogram bin and the corresponding histogram height is  $h_t$ . Bin range or width is an interval  $b$ , which is normally equal in each histogram. The number of bins is

$$n = \frac{\max(a_{ij}) - \min(a_{ij})}{b}$$

Histograms  $H_{a_{ij}}$  are used to compute the probability of the measurement at each cell.

### 3. Indoor localization using deterministic methods:

In deterministic approach, signal strength of an access point at a location is represented by a scalar value, for example, the mean, median, mode value.

In the location estimation phase the measurements are compared to the radio map to compute the location estimate. In the deterministic location estimation [4] the state  $x$  is not considered as a continuous random variable. The objective is not to obtain the whole distribution of the state  $x$  which is the RSS but to compute a single location estimate at every time step. Deterministic location estimation is based on the similarity of the measurement  $y$  which is The RSS from access point and the fingerprints  $R_i$ .

In this section the mathematical formulation of the deterministic approach is presented, the different distance measures are described and they are used in the K-nearest neighbor and in the weighted K-nearest method.

There are different measurements that are used to find the best match between observations and the radio map. The common choice for the comparison measure is to use the p-norm to assign a non-negative value to the fingerprint  $R_i$  [4].

Let  $p \geq 1$  be a real number. Then the  $p$ -norm of the  $n$ -dimensional vector  $x$  is given by [4]:

$$\|X\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p} \tag{7}$$

Substituting  $x_i = \bar{y}_i - \bar{a}_{ij}$ , where  $\bar{y}_i$  is the average of RSS values measured from access point AP, in Eq. (7) gives:

$$\|X\|_p = (\sum_{i=1}^n |\bar{y}_i - \bar{a}_{ij}|^p)^{1/p} \tag{8}$$

The  $p$ -norm  $\|X\|_p$  is the distance in the signal strength space between the measurement  $\bar{y}$  and the fingerprint  $R_i$  computed at the location estimation phase of the algorithm. Varying the parameter  $p$  leads to different norm values. For example  $p = 1$  corresponds to Manhattan distance and  $p = 2$  implies Euclidean norm, which is the most commonly used distance measure in the location fingerprinting.

### 3.1 K-nearest neighbor

K-nearest neighbor (KNN) method is one of the simplest ways to determine the location of the MU by using the radio map. KNN algorithm is a location fingerprinting method that considers K CPs to calculate the approximate position of the user. The idea is to compare the fingerprints in the radio map to the observed measurements and to select K calibration points with the "nearest" RSS values.

In the KNN approach, the vector  $\bar{y}$  is used as a measurement and compared to the radio map. Compute the distance and sort it up, Sort neighbors and select RPs whose  $D_i$  is lower. The Euclidean norm is widely used, but the Manhattan norm is also common. The most common choice as a MU's location estimator  $\hat{x}$  is the average of the coordinates of the K-nearest neighbors fingerprints, that is

$$\hat{x} = \frac{1}{K} \sum_{i=1}^k p_i \tag{9}$$

Where  $p_i$  is the position of the nearest neighbors.

### 3.2 Weighted K-nearest neighbor

One approach is to calculate the location of the MU as a weighted average of the fingerprint locations, that is

$$\hat{x} = \frac{1}{\sum_{i=1}^k w_j} \sum_{i=1}^k w_i p_i^k, \quad p_i^k \in L_K^n \tag{10}$$

Li et al. [2006] [6] use the inverse of the RSS distance as a weight, that is

$$w_i = d(\bar{y}, \bar{a}_i)^{-1},$$

Li et al. [2006] [6] assert that in general the KNN and the weighted K-nearest neighbor (WKNN) can achieve better accuracy than NN method, particularly with parameter values  $K = 2, 3$  and  $K = 4$ . However, if the density of the radio map is high, NN method can perform as well as the more complicated methods Li et al. [2006][6].

#### 4. Indoor localization using probabilistic methods:

Probabilistic approach [5] exploits the sample of measurements collected during the calibration. In this section the static location estimation is considered, the distribution of the state is computed from the simultaneously taken measurements, which are independent of the previous and the future measurements or location estimates. The idea in the probabilistic methods in location fingerprinting is to compute likelihood function.

Using Baye's theorem, this can be rewritten as:

$$\text{argmax}_i [P(i/y)] = \frac{\text{argmax}_i P(y/i).P(i)}{P(y)} \quad (11)$$

Where  $y$  is RSS values measured from access point AP and  $i$  represent the calibration point.

Since  $P(y)$  is constant for all  $i$ , [7] we can rewrite equation

$$\text{argmax}_i [P(i/y)] = \text{argmax}_x [P(y/i).P(i)] \quad (12)$$

$P(i)$  can be determined from the user profile based on the fact that if the user is at a given location, it is more probable that he will be at an adjacent location in the future. If the user profile information is not known, or not used, then we can assume that all the locations are equally likely and the term  $P(i)$  can be factored out from the maximization process

$$\text{argmax}_x [P(i/y)] = \text{argmax}_i [P(y/i)] \quad (13)$$

The term  $[P(y/i)]$  is called the likelihood function and there are several approaches for computing the likelihood function with (non-parametric and parametric Estimation).

where  $p(y|i) = p v_i(y - \bar{a}_i)$  and  $v_i = y - \bar{a}_i$ . We assume that the components of the random vector  $v_i$  are independent. Thus,

$$P(y/i) = \prod_j P v_{ij} (y_j - \bar{a}_{ij}) \quad (14)$$

##### 4.1 Parametric Approach

In the parametric approach [8], the RSS histogram is approximated as a known function. In the parametric approach, the RSS histogram is approximated as a known function. Fingerprints include a lot of data, and thus it would be practical to use some parametric methods to approximate the histograms. Unfortunately, the distribution of the RSS varies as a function of location and time because of the complex radio signal propagation environment. RSS histograms can be asymmetric and multimodal. Thus, the modeling of the measurement error distribution with a known function at every location becomes challenging. In this paper, the left-skewness of some of the histograms is exploited by approximating the histograms with a log-normal distribution [12], when the absolute values of the signal strengths are used. The left-skewness occurs due to the observation that the variations of the weaker RSS values are larger than the stronger RSS values.

The Gaussian approximation of the histograms has been widely discussed in the literature [12, 13]. The Gaussian approximation can provide a good fit with some of

the histograms, but in the case of skewed histograms the fit can be bad. Moreover, some of the histograms are quite concentrated around their mean, and the Gaussian approximation tends to spread the distribution too much. In those cases, the exponential function provides better approximation and improves there solution of different fingerprints close to each other.

There are several approaches for computing the likelihood  $p(y|i)$ . Examples of these methods are given in Table (1)

Where  $\bar{a}_{ij}$  is an average RSS values measured from access point  $AP_j$  at  $i_{th}$  calibration point,  $\sigma_{ij}^2$  is the variance of RSS values,  $y_j$  is RSS measured from access point  $AP_j$  at online phase and  $p(y|i) = pv_{ij}(y_j - \bar{a}_{ij})$  and  $v_{ij} = y_j - \bar{a}_{ij}$ .

### Non-Parametric Approach

#### 4.2.1 Kernel Method

The kernel method is a non-parametric method, because the idea is to estimate the underlying pdf from the sample pattern. The idea is to impose a probability mass to a ‘kernel’ around each of the samples  $a_{ij}^t$ .

The kernel density estimation makes it possible to interpolate the data to the entire signal strength space and fill the possible incorrect gaps in the RSS histograms. The kernel density estimation can be considered as a more general density estimation technique. [9]

(Kernel function) A kernel function is a non-negative integrable function satisfying the conditions

$$K(x) = K(-x) \tag{18}$$

$$\int K(x) dx = 1 \tag{19}$$

The computation of the likelihood is done by using the equation

$$P(y|i) = \frac{1}{|a_{ij}|h} \sum_1^{|a_{ij}|} K\left(\frac{y-a_{ij}}{h}\right) \tag{20}$$

where  $K(\cdot)$  denotes the kernel function,  $a_{ij}^t$  is the  $t_{th}$  element of the  $a_{ij}$  vector and  $h > 0$  is a smoothing parameter, which determines the width of the kernel [11]. Some of these kernel functions are:

#### Gaussian kernel

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{x^2}{2}\right)} \tag{21}$$

#### Exponential kernel

$$K(x) = \frac{1}{\sqrt{2}} e^{(-|x|)} \tag{22}$$

Where  $x = \frac{y-a_{ij}}{h}$

### 4-3 Histogram Comparison

The likelihood function is computed by using directly the whole histograms. The fingerprint histograms  $Ha_{ij}$  and the measurement histograms from the location estimation phase  $H_y$  are compared and the likelihood is computed from the similarity value between these histograms.

Histogram comparison

$$P(y/i) = d_{pdf}(H_{y_j}, H_{a_{ij}}). \quad (23)$$

Different probability density distance measures  $d_{pdf}(\cdot, \cdot)$  for the histogram comparison method are given in table (2)[1]

## 5. Experimental Results

Using personal laptop With Vistumbler program that retrieves the basic information of each Aps (SSID, RSS) and RSS values are in dBm scale.

The radio map were collected in a department corridor with a dimension  $8 \times 85 \text{ m}^2$  which divided into 20 grid cells and the CPs at the center of the cells as shown in **Error! Reference source not found.** . RSS values from one AP at the CPs varied through the grids due to obstacles such as human, doors...etc., all the grids has the same area.

The radio map was collected at 20 CPs with a fixed orientation but different RSS samples at different time 1 sample/sec.

Three Radio maps were constructed ,The first radio map has 60 RSS samples with 60sec measurement period, the second has 100 RSS samples with 100 sec measurement period and the second has 300 RSS samples with 300 sec measurement period using personal laptop .

With Vistumbler program that retrieves the basic information of each Aps (SSID, RSS) and RSS values are in dBm scale.

There are 3 Aps were placed inside the corridor so that all Aps are heard at each CPs

Note that the samples of default reading more than -90dBm and lower than -29dBm are excluded.

The K-nearest neighbor method and its modifications were tested. Different numbers of tested points were considered in the computation of the estimate by varying the parameter K (k=2, k=3, k=4 ) and the location estimate was computed as the average of the coordinates of the K'' nearest'' CPs. Different measures were used to compute the distances in the signal space. The effect of varying the parameter K in KNN method is examined and locating the stationary MU is considered first. The Radio maps were constructed as discussed in section 2.

In this section the static location estimation algorithms studied in this work are examined. Each of the location estimation algorithms has several parameters which were varied to find the best performance. The root mean square error (RMSE) was used as the primary performance measure to obtain the best algorithm.



The p-norm used to measure the distance by varying parameter p one gets different norms p = 1 corresponds to Manhattan distance, p = 2 implies Euclidean norm and p=3, p=4 and p=5.

The KNN algorithm and its modifications were applied on the three Radio maps to obtain the suitable sample size of the Radio map for the experiments. There are four chosen stationary test cases, test points 1, 2, 3 and 4 and many different types of cases for example at Radio map 60 samples using p = 2 implies Euclidean norm and RMSE to obtain which Raw data that contain average, median and mode is the best and which one of the parameter K (k=2, k=3, k=4 ) is suitable for this experiment. Figure 2 show this case with 3 APs.

In table 3 the effect of varying the parameter k (k=2,3,4) at p-norm p=1, Radio map 60 samples with different raw data. Using RMSE as performance measure.

In table 4 the effect of varying the parameter k (k=2,3,4) at p-norm p=1, Radio map 60 samples with different raw data. Using RMSE as performance measure.

The Results after all the experiments with 3 APs by varying all the parameters P, K, raw data and raw map samples obtained the best estimated position for KNN algorithm at raw map 100 samples with mode raw data at P=2 Euclidian distance and K= 2 that gives the smallest value of MSE 0.3633.

And the best estimated position for WKNN algorithm at raw map 100 samples with mode raw data at P=5 and K= 2 that gives the smallest value of MSE 0.3086.

So that the WKNN algorithm is better than the KNN algorithm at Deterministic location estimation. Figure (3, 4) shows the Effects of the parameters on performance such as No of Aps in KNN Algorithm and Figure (5, 6) shows the Effects of the parameters on performance such as No of Aps in WKNN Algorithm .

As the result of all experiments obtained that using three Aps is more suitable than each other due to the MU heard the 3 Aps or at least one so the error was smaller than each one.

Histogram comparison method: The Bhattacharyya distances provided the smallest ME at raw map data 100 samples as shown in Table (5). .

Parametric approximation: The measurement likelihood was computed using different parametric approximations of the normalized histograms. In the tests, the exponential and Gaussian approximation produced the smallest RMSE as shown in Table (6).

Kernel method: The kernel method was tested with several kernel functions and kernel widths. Different kernel functions were tested with kernel width  $h = 3$  which provided the smallest RMSE as shown in Figure (5). The exponential kernel function is claimed to provide good results at raw map 100 samples and in this work, it also gave better performance compared to the another kernel function as shown in Table(7).

**Conclusion**

In this paper, different location fingerprinting methods were considered by introducing the mathematical basis of the methods and testing them with WLAN RSS measurement data. The mathematical formulation was carried out from different points of view, and the parameters of the methods were varied in the tests in order to obtain the best performance. The environment variables, such as the number of access points (APs) and the radio map density, were also varied and the methods were compared also in these varying circumstances. The main goal in the mathematical formulation was to model the location as a random variable which led to the division of the area of interest into rectangular cells instead of considering only the individual calibration points. Deterministic methods, such as the K-nearest neighbor (KNN) method and the Weighted K-nearest neighbor(WKNN) method, were formulated with the discrete location variable, as they are presented in the literature .probabilistic methods were formulated by estimation of likelihood functions with several approaches (non-parametric and parametric) are also explained and can be computed from the similarity value between Histogram (Histogram comparison) .

Table 1: The approaches for computing the likelihood p(y|i) in the parametric approach

Gaussian Method	$Pv_{ij}(y_j - \bar{a}_{ij}) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} e^{-\frac{(y_j - \bar{a}_{ij})^2}{\sigma_{ij}^2}} \quad (15)$
Exponential Method	$Pv_{ij}(y_j - \bar{a}_{ij}) = \frac{1}{2} e^{-(y_j - \bar{a}_{ij})} \quad (16)$
Log-normal Method	$Pv_{ij}(y_j - \bar{a}_{ij}) = \frac{1}{(y_j - \bar{a}_{ij})\sqrt{2\pi\sigma_{ij}^2}} e^{-\frac{(\log(y_j - \bar{a}_{ij}) - \bar{a}_{ij})^2}{\sigma_{ij}^2}} \quad (17)$

Table 2: PROBABILITY DENSITY DISTANCES

Name	$d_{pdf}(p, q)$
Bhattacharyya	$D = -\ln(\sum_1^n \sqrt{p_i * q_i})$
Squared chord distance	$D = \sum_1^n (\sqrt{p_i} - \sqrt{q_i})^2$

Kullback-lubler distance	$D = \sum_1^n p_i * \ln(\frac{p_i}{q_i})$
Canbera distance	$D = \sum_1^n (\frac{ p_i - q_i }{p_i + q_i})$
Lorentzian distance	$D = \sum_1^n \ln(1 +   p_i - q_i  )$

Table 3: RMSE as performance measure

At raw data 60 samples	K=2 RMSE (m)	K=3 RMSE (m)	K=4 RMSE (m)
Average raw data	2.25	1.77	2.43
Median raw data	3.51	2.65	3.064
Mode raw data	3.51	3.49	3.064

Table 4: RMSE as performance measure

At raw data 60 samples	K=2 RMSE (m)	K=3 RMSE (m)	K=4 RMSE (m)
Average raw data	2.019	2.36	2.119
Median raw data	3.71	3.49	2.73
Mode raw data	4.32	3.49	3.537

Table 5: The Bhattacharyya distances provided the smallest ME at raw map data 100 samples.

Distance	At raw map 60 samples RMSE (m)	At raw map 100 samples RMSE (m)	At raw map 300 samples RMSE (m)
Bhattacharyya	3.95	3.68	3.86
Squared chord distance	4.62	5.34	6.27
Kullback-lubler distance	5.65	5.86	5.87
Canbera distance	6.32	5.32	4.23
Lorentzian distance	4.369	6.325	7.13

**Table 6: The exponential and Gaussian approximation produced the smallest RMSE**

Function	At raw map 60 samples RMSE (m)	At raw map 100 samples RMSE (m)	At raw map 300 samples RMSE (m)
Gaussian Method	3.48	3.0822	2.9155
Exponential Method	3.8079	3.08	2.9155
Log normal Method	4.023	3.969	3.632

**Table 7: The exponential kernel function is claimed to provide good results at raw map 100 samples**

Function	At raw map 60 samples RMSE (m)	At raw map 100 samples RMSE (m)	At raw map 300 samples RMSE (m)
Gaussian kernel	3.632	3.356	3.21
Exponential kernel	3.35	3.032	3.156

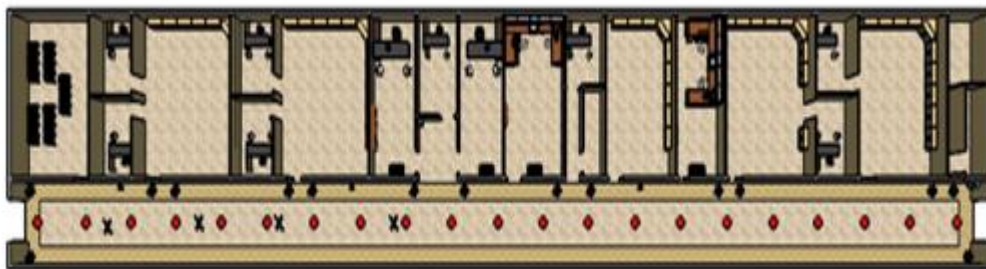


Figure 1: The distribution of APs and CPs inside the corridor

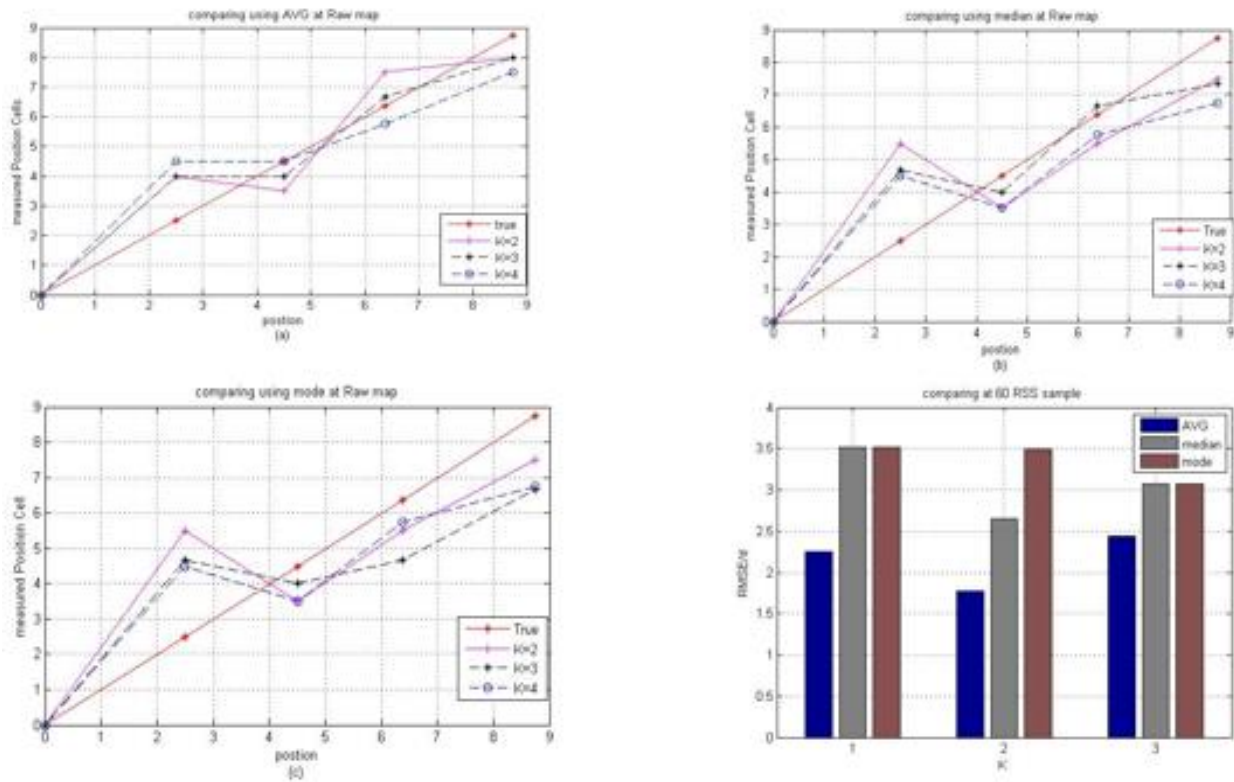


Figure 2: a) The deviation the estimated positions at AVG raw data. b) Median raw data c) mode raw data. d) Show the RMSE due to varying K parameters and raw data.

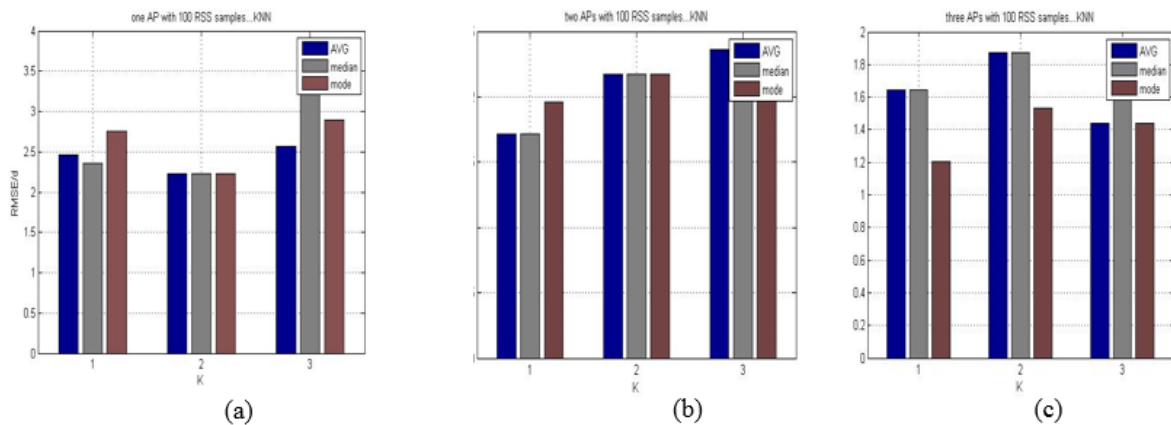


Figure 3: Using 3 Aps is more suitable than one or two Aps and this decision was based on the least error at the same parameters, for KNN algorithm at raw map 100 samples P=2 Euclidian distance which gave the best results at KKN.

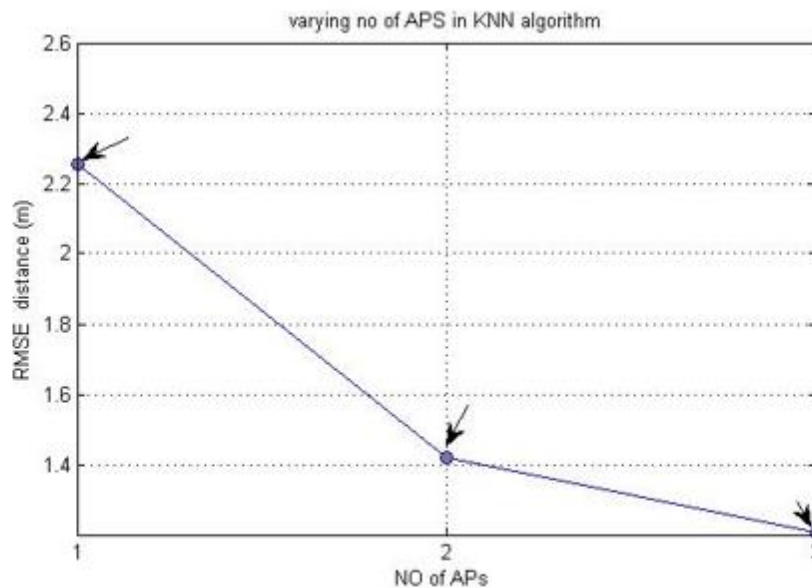


Figure 4: comparing between using one, two or three Access points APs

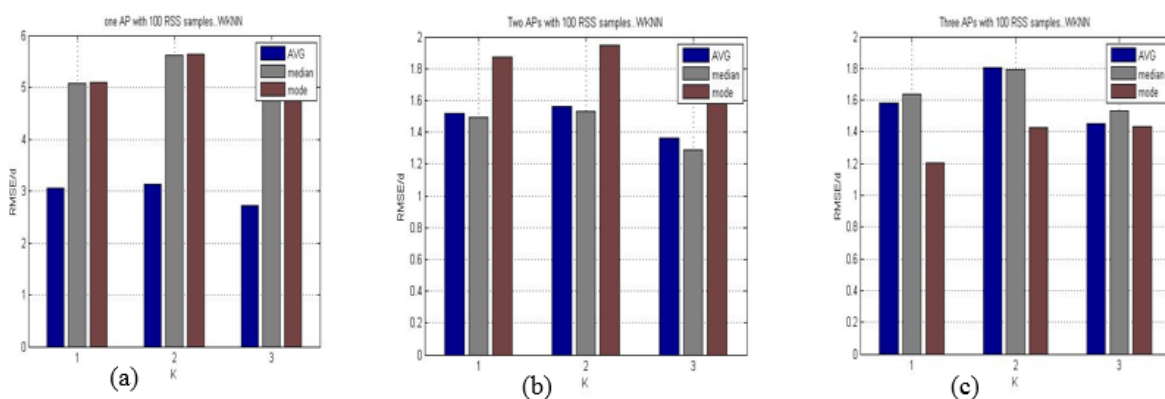


Figure 5: Using 3 Aps is more suitable than one or two Aps and this decision was based on the least error at the same parameters, for WKNN algorithm at raw map 100 samples at P=5 which gave the best results at WKKN.

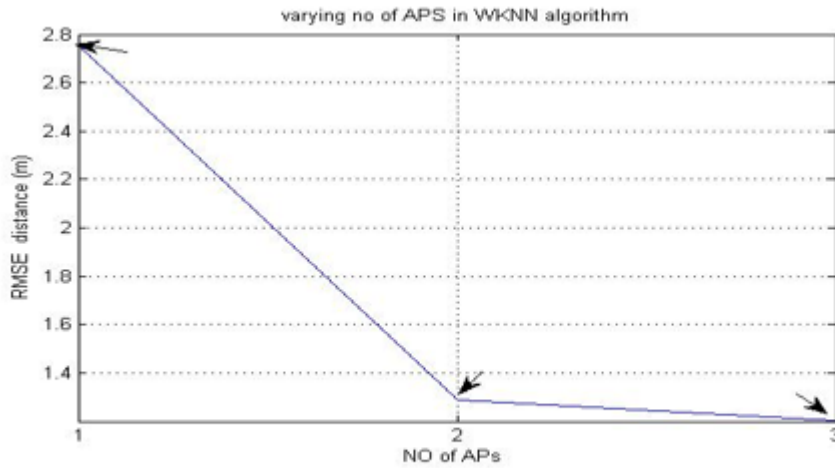


Figure 6: comparing between using one, two or three Access points Aps

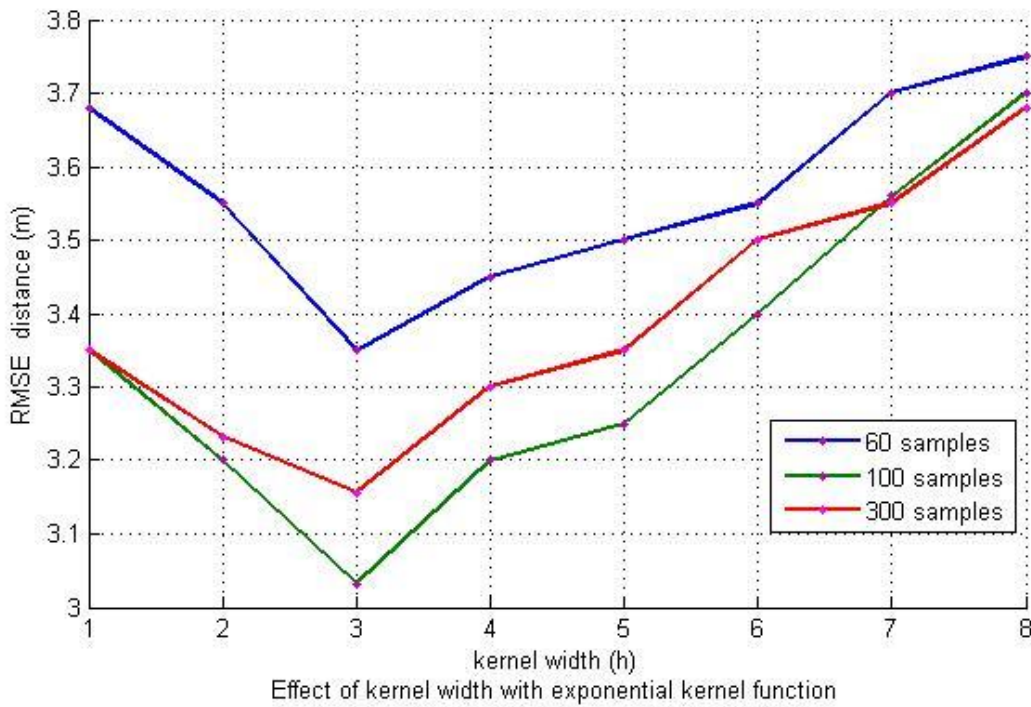


Figure 5: The kernel width  $h = 3$  which provided the smallest RMSE

## References

- [1] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche, "A comparative survey of WLAN location fingerprinting methods," in Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on, March 2009, pp. 243–251.
- [2] B. Li, J. Salter, A. G. Dempster, and C. Rizos, "Indoor positioning techniques based on wireless lan," in 1st IEEE Internal Conference on Wireless Broadband & Ultra Wideband Communications, March 2006, pp. 13–16.
- [3] ParamvirBahl and Venkata N. Padmanabhan. Radar: An in-building rf-based user location and tracking system. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2(10):775–784, March 2000.
- [4] ParamvirBahl and Venkata N. Padmanabhan. Radar: An in-building rf-based user location and tracking system. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2(10): 775–784, March 2000.
- [5] TeemuRoos, Petri Myllymäki, Henry Tirri, Pauli Misikangas, and JuhaSievänen. A probabilistic approach to wlan user location estimation. International Journal of Wireless Information Networks, 9(3):155–163, July 2002.
- [6] Binghao Li, James Salter, Adrew G. Dempster, and Chris Rizos. Indoor positioning techniques based on wireless lan. Technical report, School of Surveying and Spatial Information Systems, UNSW, Sydney, Australia, 2006.
- [7] Rehim, A. A. Y. A. and M. Amin (2004). HORUS: A WLAN-based indoor location determination system.
- [8] Honkavirta, V., et al. (2008). "Location fingerprinting methods in wireless local area networks." Master of Science Thesis, Tampere University of Technology, Finland.
- [9] Honkavirta, V., et al. (2009). A comparative survey of WLAN location fingerprinting methods. Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on, IEEE.
- [10] Cha, S.-H. (2007). "Comprehensive survey on distance/similarity measures between probability density functions." City 1(2): 1.
- [11] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification. John Wiley, Sons Inc., 2001.



- [12] K. Kaemarungsi, "Distribution of WLAN received signal strength indication for indoor location determination," National Electronics and Computer Technology Center, Thailand, Tech. Rep., 2006.
- [13] B. Li, J. Salter, A. G. Dempster, and C. Rizos, "Indoor positioning techniques based on wireless LAN," School of Surveying and Spatial Information Systems, UNSW, Sydney, Australia, Tech. Rep., 2006.