



# Improving Positioning Accuracy Using WLAN Optimization for Location Based Services and Cognitive Radio Networks

Sohaib Bin Altaf Khattak<sup>1</sup>, Min Jia<sup>1,2(✉)</sup>, Qing Guo<sup>1</sup>, and Xuemai Gu<sup>1</sup>

<sup>1</sup> Communication Research Center, School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China  
{sohaib, jiamin, qguo, guxuemaii}@hit.edu.cn

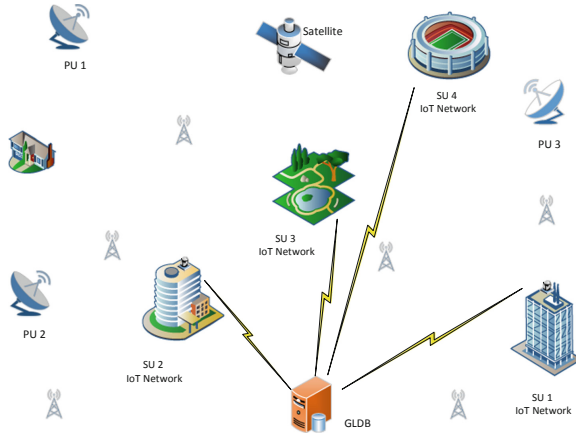
<sup>2</sup> CETC Key Laboratory of Aerospace Information Applications, Shijiazhuang, China

**Abstract.** Positioning information not only benefits the localization systems but also improve the performance of geo-location based cognitive radio (CR) networks. Most researchers focus on other aspects of CR and databases but rarely discuss the fact of how the positioning information can influence the performance of CR systems in indoor environments. WLAN is the most common technology used for indoor positioning. Optimization of WLAN access points (APs) can enhance accuracy of the localization systems. In this paper, we present an optimization algorithm for WLAN localization system. The proposed scheme estimates the optimal density of the APs required to meet the coverage demands and optimize their deployment to enhance the localization accuracy. One of our main contributions is the APs hearability-based reference points (RPs) clustering technique. Its uniqueness lies in the fact that not all installed APs participate in the localization process for all RPs. Finally, we analyze the variables governing the optimization process and the trade-off between cost, computation, and accuracy. Extensive simulations are conducted to validate the effectiveness of our algorithm. Our approach reduces the mean positioning error by 25% and the maximum error by 44% compared to the previous algorithm's performance.

**Keywords:** Access point · Cognitive radios · Geo-location database · Indoor localization · Location based services · WLAN optimization

## 1 Introduction

Location-Based Services (LBS) are one of the most facilitative and beneficial applications. These applications demand location information of the connected devices for their execution [1]. Received data without location information is of no or limited use. These connected smart devices continuously send and receive data, which is utilized for a dynamic range of applications [2]. This massive amount of data generated poses a significant challenge on the available spectrum



**Fig. 1.** Cognitive radio urban scenario.

and leads towards spectrum shortage. The ISM band is the primary source of communication in indoor environments. Heterogeneous devices use the same ISM band, which makes it overcrowded. Effective sharing techniques are required to fulfil this demand [3, 4]. Instead of using the ISM band, the concept of Cognitive Radios (CR) can be applied, where the CR enabled devices to interact with each other, use a licensed spectrum when it is vacant (Fig. 1).

To find the available spectrum the Geo-location database can be used [5]. In this concept, the secondary user (SU) network is informed about the available spectrum by a spectrum database. The Primary User (PU) provide their availability information to the database location wise. On the other hand, when the SU requires to access the vacant spectrum, they also need to provide their location information. The GPS can provide reliable location information of the SUs, but when it comes to indoor environments, the GPS is ineffective. If a user or a device moves from outdoor to indoor, it must switch between GPS to an indoor positioning system (IPS). The inaccurate location information delivered to the database can cause interference between PU and SU, and also affect the utilization of the spectrum. Thus, a proper IPS needs to be connected with the CR network. Traditionally, IPS is used to locate the objects for LBS only. This location information can be used for many other services also. We aim here is to develop an efficient spectrum sharing scheme, utilizing the location information obtained from IPS.

The WLAN fingerprinting based IPS has emerged as the most popular technology [6, 7]. WLAN installation requires careful design and planning. The planning and optimization of wireless networks is an important research issue, [9, 10]. Much work has been done recently to guarantee the quality of service and network performance, [8, 11], however, the optimal design of WLAN for IPS is still an open issue. Interestingly, just the APs optimization can lead to performance enhancement in multiple ways. Therefore, we consider AP optimization as an

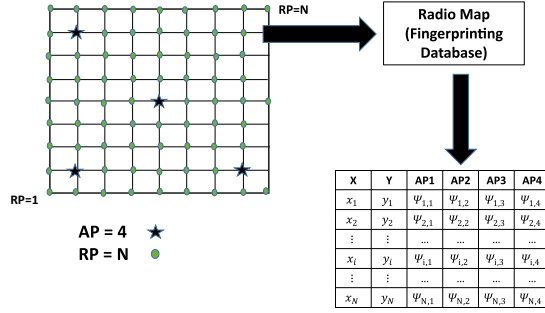


Fig. 2. Radio map for fingerprinting based indoor positioning systems.

essential topic to investigate. In our earlier work [12], we have already considered this problem for IPS. In this article, we enhance our previous work by determining the optimal APs density and improve AP deployment scheme. We test our algorithm in a more realistic and complex indoor environment for fingerprinting based positioning systems and compare it with our previous work. For ease, we call our previous algorithm as A and the enhanced algorithm as B.

## 2 System Model

We consider a 2D indoor environment as our area of interest (AoI) and divide it into equal spacing grids or RPs. The readings from APs are taken at these RPs and further used in optimization and positioning processes. We assume our AoI is big enough that a single AP can not provide coverage to all RPs. Multiple APs are installed to provide complete coverage. Thus, the adequate number of APs and their appropriate locations are vital factors to be investigated. Our algorithm will address this issue, ensuring positioning accuracy and signal coverage.

WLAN fingerprinting has two phases, the offline phase, and an online phase. In the offline phase a radiomap is constructed by collecting the RSS readings at all RPs. Each RP has a unique signature, based on the collected RSS values from the APs. A database is created using these unique signatures or fingerprints (FPs) corresponding to their RPs. A simple illustration can be seen in Fig. 2. In the online phase, a device sends a query to the database containing the unique FP it has recorded. The database matches this query with the available data and the best match is given as the estimated position of the device. Mathematically it can be represented as:

$$RP_j = \{(x_j, y_j) | j = (1, \dots, N)\} \tag{1}$$

$$\lambda = \begin{bmatrix} x_i, y_i & (\psi_{i,1}, \dots, \psi_{i,l}) \\ \vdots & \vdots \\ x_N, y_N & (\psi_{N,1}, \dots, \psi_{N,l}) \end{bmatrix} \tag{2}$$

here  $(\lambda)$  is the FP database,  $\psi =$  RSS samples from APs,  $(x, y)$  are the coordinate points of the RPs,  $N$  is the total number of RPs, and  $l$  represents the total number of APs. To simulate the indoor signal propagation, we use log-normal path loss model.

$$P(d) = P(d_0) + 10.n.\log(d/d_0) + \zeta + \sum_{i=1}^j(\gamma_i) + \sum_{k=1}^l(\rho_k) \quad (3)$$

here,  $P(d)$  is the RSS at a point  $d$  meters away from AP,  $P(d_0)$  is the RSS at reference distance (1 m),  $n$  is the path loss coefficient ( $n = 3$ ); noise ( $\zeta$ ), Wall attenuation factor ( $\gamma = 4$ ), and people attenuation factor ( $\rho = 3$ ) are the other attenuating factors.

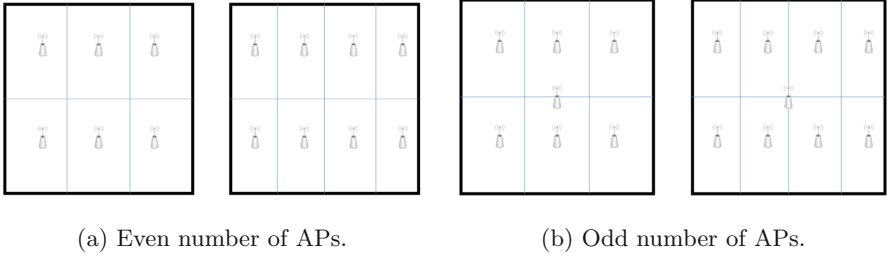
### 3 Proposed Technique

Installing a large number of APs without proper planning can lead towards performance degradation, and an increase in both cost and computation. We develop a scheme to provide the optimal number of APs, satisfying the need for coverage and accuracy.

The RPs are used for FP collection and also serve as candidate positions for APs. The first AP (AP1) is placed at its desired position by the user. RSS values at each RP are calculated, and RPs with RSS greater than  $-80$  dBm are checked. RPs satisfying the RSS condition are the average number of RPs an AP can cover. This simple calculation can give us an idea about the total number of APs required to cover all RPs. We can call this number as min APs, it is tentative and only from the viewpoint of coverage. Now, we initialize two variables,  $Ck_A$  and  $Ck_B$ .  $Ck_A$  must reach the count equal to the total number of RPs.  $Ck_B$  is the minimum number of APs that must provide coverage to each RP can be adjusted according to the requirement.

If the min APs are even in number, these APs are deployed in the AoI using symmetrical deployment, as shown in Fig. 3(a), if odd, symmetry is shown in Fig. 3(b). The basic idea of this symmetrical deployment is to divide the area into equal regions and installing each AP at the center points. After deployment, the RSS for these APs is calculated again, and if  $Ck_A$  equals  $N$ , we stop and display the number of required APs by adding one more AP. If the  $Ck_A$  condition is not satisfied, the process is repeated by increasing the min APs by one and continue the process until  $Ck_A$  condition is fulfilled. This optimal density of APs will provide coverage to every RP in AoI, with each RP having at least  $Ck_B$  APs with sufficient signal coverage. This methodology not only ensures enough number of APs for good signal coverage but also provide adequate overlapping regions for positioning purpose. As all RPs will have at least  $Ck_B$  APs to record RSS. The pseudo-code of the process can be seen in Algorithm 1.

AP deployment algorithm also starts by placing AP1 at the desired position, RSS is measured at all RPs and their average value is compared with a threshold  $Th$ , i.e. RSS between  $-55$  dBm and  $-60$  dBm. The reason for using this specific



**Fig. 3.** Symmetrical deployment of APs to analyze signal coverage of installed number of APs.

threshold window has been explained in detail in our previous paper [12]. It helps create appropriate overlapping regions among APs; without a proper overlap, signal coverage can be ensured but not the positioning accuracy. Mathematically we represent the resultant RSS readings from AP1 as  $J = [\psi_1, \dots, \psi_N]$ , where  $N$  is the total number of RPs. By applying the threshold condition,  $J$  is converted into a matrix of 0s and 1s:

$$\psi_i = \begin{cases} 1, & \text{if } -55 \geq \psi_i \geq -60 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Those RPs satisfying Eq. (4) are taken as candidate positions for next AP, and saved in  $D(1)$ .

$$D(1) = \begin{bmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_l & y_l \end{bmatrix} \quad (5)$$

$l$  is the total number of new candidate locations, and  $(x, y)$  are the coordinates. The same process is applied for all points in  $D(1)$ , and the new points are saved in  $D(2)$  and so on. Let us suppose the matrix  $D(1)$  had four points, so  $D(2)$  will have four arrays, each corresponding to the point in  $D(1)$ .

$$D(2) = \begin{bmatrix} x_{11} & y_{11} \\ \vdots & \vdots \\ x_{1l_1} & y_{1l_1} \end{bmatrix} \begin{bmatrix} x_{21} & y_{21} \\ \vdots & \vdots \\ x_{2l_2} & y_{2l_2} \end{bmatrix} \begin{bmatrix} x_{31} & y_{31} \\ \vdots & \vdots \\ x_{3l_3} & y_{3l_3} \end{bmatrix} \begin{bmatrix} x_{41} & y_{41} \\ \vdots & \vdots \\ x_{4l_4} & y_{4l_4} \end{bmatrix} \quad (6)$$

Here,  $l_1$  to  $l_4$  are the lengths of the four arrays independent of each other. After making a database of candidate positions, the next step is to make all possible APs configurations; the configurations are extracted from this database as a tree structure, explained in our previous work [12]. The updated algorithm is shown in Algorithm 2.

**Algorithm 1.** Optimal Density of Access points.**Input:** AP1 =  $(x, y)$ **Output:** Total APs =  $T_{AP}$ 


---

```

1: Calculate  $RSS$  at RPs from AP1
2: Match  $RSS$  with  $Th$  of  $-80$  dBm
3: No. of RPs satisfying condition (2) = avg. pts.
4:  $minAPs$  = total RPs/avg. pts
5:  $T_{AP}=minAPs$ 
6:  $Ck_A=0$  ( $Ck_A < lc$ )  $lc$ = total no. of RPs
7: Check  $minAPs$ 
8: if  $minAPs$  is even then
9:   Plot APs in symmetry 1
10: else  $\{N$  is odd $\}$ 
11:   Plot APs in symmetry 2
12: end if
    LOOP Process
13: for  $i = 1$  to  $T_{AP}$  do
14:   Calculate  $RSS$ 
15:   Match  $RSS$ ,  $Ck_B$  APs must provide coverage above  $Th2 = -100$ 
16:   if  $Ck_A=lc$  then
17:      $T_{AP}=T_{AP} + 1$ 
     STOP:  $T_{AP}$  is optimal
18:   else
19:      $T_{AP}=T_{AP} + 1$ 
     Go back to Step 7
20:   end if
21: end for
    End LOOP Process
    End While
22: return  $T_{AP}$ 

```

---

The used hybrid technique is a combination of *MaxFD* and *MinGDOP*, i.e., maximizing fingerprint difference (FD) and minimizing Geometric Dilution of Precision (GDOP). *MaxFD* and *MinGDOP* have values in different ranges. We normalize it on the scale of 0 to 1.

$$MaxFD = argmax \sum_{i=1}^N \sum_{j=1}^N d(\delta_i, \delta_j) \quad (7)$$

$$MinGDOP = argmin \sqrt{Trace(H)} \quad (8)$$

$$Hybrid = argmax \left( MaxFD + \frac{1}{MinGDOP} \right) \quad (9)$$

We use a cluster-based approach for selection strategy and calculation of objective function. The Hamming distance is used to divide the data into clusters. Before applying Hamming distance, the data is transformed into a binary

---

**Algorithm 2.** Access Points Deployment strategy.

---

**Input:** AP1 =  $(x, y)$  & Total APs =  $T_{AP}$

**Output:** APs Configurations

```

LOOP Process
1: for  $i = 1$  to  $T_{AP}$  do
2:   Calculate  $RSS$  at RPs from  $AP(i)$ 
3:   Match  $RSS$  with  $Th$  from Eq. (4)
4:   Save the candidate positions in  $D\{i\}$ 
5:    $AP\{i\} = D\{i\}$ 
6: end for
End LOOP Process
7:  $l=1$ 
LOOP Process
8: for  $i = 1$  to  $length(D)$  do
9:   for  $j = 1$  to  $length(D\{i\})$  do
10:    for  $k = 1$  to  $size(D\{i\}\{j\}, 1)$  do
11:      Configuration( $l$ )=  $[AP1\ D\{i\}\{j\}\ D\{i\}\{j\}\{k\}]$ 
12:       $l=l+1$ 
13:    end for
14:  end for
15: end for
End LOOP Process
16: Configurations
17: DELETE configurations having repetition
18: Display (Configurations)

```

---

representation. Hamming distance can be defined as the number of bits between two bit sequences in which it differs. Having two bit sequences  $A_i$  and  $B_i$ , mathematically it can be represented as:  $HD(A_i, B_i) = q + r$ . Here,  $q$  represents the number of bits with value 1 in  $A_i$  and 0 in  $B_i$ , and  $r$  is vice versa. And also  $HD(A_i, B_i) = HD(B_i, A_i)$ .

Different APs have different positioning abilities; each AP contributes differently to the positioning process. Similarly, in an area where a single AP cannot provide full coverage, different sets of APs are visible at different RPs. Some sets can contribute more to the positioning process than others, based on the number of RPs covered. As an example we consider  $T_{AP_s}$  as 7 in AoI  $X$ . The RSS data can be represented as:

$$\Theta = \begin{bmatrix} (\psi_{1,1}, \psi_{1,2}, \psi_{1,3}, \psi_{1,4}, \psi_{1,5}, \psi_{1,6}, \psi_{1,7}) \\ \vdots \\ (\psi_{i,1}, \psi_{i,2}, \psi_{i,3}, \psi_{i,4}, \psi_{i,5}, \psi_{i,6}, \psi_{i,7}) \\ \vdots \\ (\psi_{N,1}, \psi_{N,2}, \psi_{N,3}, \psi_{N,4}, \psi_{N,5}, \psi_{N,6}, \psi_{N,7}) \end{bmatrix} \tag{10}$$

For simplicity we rewrite the matrix in row form and replace  $(\psi_{i,1}, \dots, \psi_{N,7})$  for  $RP_i$  by  $\delta_i$ .

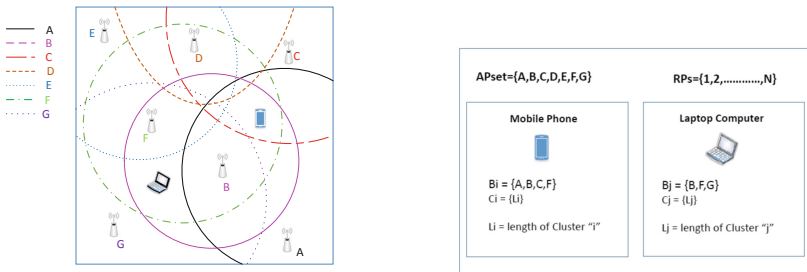
$$\Theta = [(\delta_1, \delta_2, \delta_3, \dots, \delta_N)] \tag{11}$$

Those RPs where the RSS of an AP is below  $-100$ , is considered as a dead spot for that AP. This is considered as sensitivity threshold and indicates that a particular AP is not hearable. For mathematical simplification, we replace all the RSS values less than  $-100$  dBm as  $-100$  dBm. The values in this range are practically so weak that it cannot be considered for any use, whether data transmission or positioning purposes.

At an  $RP_i$ , four APs are hearable, and three are not hearable among  $T_{APs}$ . So it can be transformed into binary by representing the hearable APs as 1 and others as 0.

$$\Theta = \begin{bmatrix} (1, 1, 1, 0, 1, 1, 0) \\ \vdots \\ (1, 1, 0, 0, 1, 1, 0) \\ \vdots \\ (1, 1, 0, 0, 0, 0, 0) \end{bmatrix} \tag{12}$$

Suppose the dimensionality of the binary matrix is  $N \times T_{AP}$ , and is partitioned into  $K$  clusters,  $C = (C_i, \dots, C_K)$ .  $RP_a, RP_b, RP_c, RP_d$  form a cluster  $C1$ , as they share the same APs set  $B1 = [AP1, AP2, AP3]$ , covering them, while  $RP_e, RP_f, RP_g, RP_h$  form another cluster  $C2$ , as they form RPs from the set  $B2 = [AP4, AP5, AP6, AP7]$  of APs and so on. A laptop and a mobile phone can be seen in Fig. 4(a), both of these devices belong to different clusters with the visibility of different sets of APs.



(a) Deployment of APs forming clusters. (b) Two devices with coverage from different sets of APs.

**Fig. 4.** APs hearability based clustering technique.

Each cluster obtained by the algorithm must be unique and sensitive to the order of APs. The criteria of clustering must satisfy certain conditions. The set of APs and RPs is  $AP = [1, 2, \dots, T_{AP}]$  and  $RP_s = [1, 2, \dots, N]$  respectively, where  $B_i \subset AP$ . We can also say that  $B_1 \cup B_2 \cup B_3 \cup \dots \cup B_K = AP$  and  $C_i = \{RP_i | i \in (1 \dots N) \& B_i \subset AP\}$ . An RP is assigned to one and only one cluster. All clusters lengths together  $C_1 \cup C_2 \dots \cup C_n = RPs$ . The size

of two clusters  $C_i$  and  $C_j$  can or cannot be equal, depending upon the  $B_i$  and  $B_j$  covering the number of RPs. If the difference between  $B_i$  and  $B_j$  is zero, then  $C_i$  and  $C_j$  are same. If the set  $B_i$  of APs covers the majority of the RPs,  $C_i$  will have the most influence on the overall system. As shown in Fig. 4 and Fig. 5, few APs are not hearable at  $C_i$  so they do not participate in the positioning process.

Let us assume, in  $C_1$  the set of APs is  $B_1$ , and it contains 4 number of APs, and total RPs covered in  $C_1$  are  $L$ . For each cluster the FPs are calculated and stored in the database. To take the FD, we take the RSS values from Eq. (2) and rewrite it as shown in Eq. 10, and Eq. 11. The difference is measured using Euclidean distance:

$$FD(i, j) = \sqrt{(\psi_{1,i} - \psi_{1,j})^2 + \dots + (\psi_{4,i} - \psi_{4,j})^2} \quad (13)$$

$$\Delta = \begin{bmatrix} (\delta_i - \delta_i & \delta_i - \delta_{i+1} & \dots & \delta_i - \delta_L) \\ (\delta_{i+1} - \delta_i & \delta_{i+1} - \delta_{i+1} & \dots & \delta_{i+1} - \delta_L) \\ \vdots & \vdots & \vdots & \vdots \\ (\delta_L - \delta_i & \delta_L - \delta_{i+1} & \dots & \delta_L - \delta_L) \end{bmatrix} \quad (14)$$

$$\Delta = \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} & \dots & \varphi_{1,L} \\ \varphi_{2,1} & \varphi_{2,2} & \dots & \varphi_{2,L} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{L,1} & \varphi_{L,2} & \dots & \varphi_{L,L} \end{bmatrix} \quad (15)$$

$$\Omega_i = \frac{\sum_{j=1}^L \varphi_{i,j}}{L} \quad (16)$$

$$O = [(\Omega_1, \Omega_2, \Omega_3, \dots, \Omega_L)] \quad (17)$$

$$MaxFD_{C_i} = \frac{\sum_{i=1}^L \Omega_i}{L} \quad (18)$$

We now have FD  $\Omega$  for each RP; to calculate the cluster-wide FD we retake the mean value. After calculating the FD for all clusters, the next step is to assign appropriate weights and calculate the final network-wide FD for this specific configuration. A cluster covering more RPs will have more influence on overall accuracy and needs to be assigned more weight as compared to a cluster covering a limited number of RPs. This weightage is assigned to the clusters, as shown below:

$$\sum_{i=1}^K (MaxFD_{C_i} \times (\frac{L_i}{N})) \quad (19)$$

To calculate the GDOP, we adopt the same process of clusters. Let  $(x, y)$  be the coordinates of RP and  $(x_i, y_i)$  are coordinate points of the  $i_{th}$  AP. For cluster  $C_1$ , the GDOP will be calculated as shown below:

$$R_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad (20)$$

---

**Algorithm 3.** Selection strategy for optimal AP configuration.

---

**Input:** All Configurations

**Output:** Optimal Configuration

- 1: **for**  $S = 1$  to  $\text{length}(\text{Configurations})$  **do**
  - 2:   Calculate  $RSS$  at all  $RPs$  from all  $APs$
  - 3:   Make Fingerprints
  - 4:   Replace  $RSS < -100$  dBm to  $-100$  (Sensitivity Th)
  - 5:   Form Clusters on the Basis of APs hearability
  
  - 6:   **for**  $i = 1$  to  $\text{length}(\text{Cluster})$  **do**
  - 7:     Calculate FD b/w all sets of two RPs within Cluster  $C_i$
  - 8:     Cal. FD for each RP and then cluster wide FD for  $i$
  - 9:   **end for**
  - 10:   Now calculate Network Wide FD, with assigning weights
  
  - 11:   **for**  $i = 1$  to  $\text{length}(\text{Cluster})$  **do**
  - 12:     Calculate GDOP for each RP then cluster wide GDOP for  $i$
  - 13:   **end for**
  - 14:   Calculate Network Wide GDOP
  - 15: **end for**
  - 16: Normalize both calculated FD and GDOP on the scale of 0 to 1.
  - 17: The Configuration with the best Score will be the optimal configuration
  - 18: Validate using both Deterministic and Probabilistic Matching Algorithms
- 

$$A = \begin{bmatrix} \frac{x_1-x}{R_1} & \frac{y_1-y}{R_1} \\ \frac{x_2-x}{R_2} & \frac{y_1-y}{R_1} \\ \frac{x_3-x}{R_3} & \frac{y_1-y}{R_1} \\ \frac{x_4-x}{R_4} & \frac{y_1-y}{R_1} \end{bmatrix} \tag{21}$$

The matrix  $A$  is multiplied by its transpose and then its inverse is calculated. If the obtained matrix is  $H$ , we calculate the GDOP by

$$\text{MinGDOP} = \sqrt{\text{Trace}(H)} \tag{22}$$

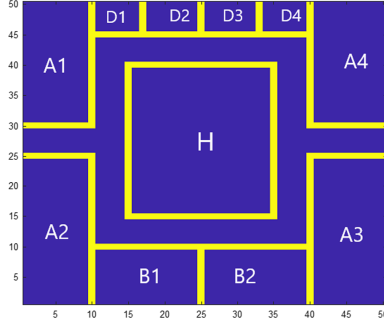
$$\text{MinGDOP}_{C_i} = \frac{\sum_{j=1}^L \text{GDOP}(i)}{L} \tag{23}$$

$$\sum_{i=1}^K (\text{MinGDOP}_{C_i} \times (\frac{L_i}{N})) \tag{24}$$

We take the normalized values of the calculated FD and GDOP. This selection scheme is summarized in Algorithm 3.

## 4 Simulation Results

We consider a 2-D area  $X$  with dimensions of  $50 \times 50$  m, as shown in Fig. 5. RPs with 2 m and 3 m spacing will be used in two different scenarios. We divide the positioning area into different rooms, and each room has people's presence.  $H$  has 40,  $A(1-4)$  has 20 each,  $B(1,2)$  has 10 each and  $D(1-4)$  has 1 each, total crowd is 144 people. This presence of people is included along with walls, obstacles, and noise in the optimization process.



**Fig. 5.** Illustration of the floor plan for our simulations in MATLAB.

To evaluate the performance of obtained AP configuration, 100 random test points are generated within the AoI. The 100 queries can be presented as  $S$  in the positioning model. KNN deterministic algorithm and Bayesian probabilistic algorithm are used to validate our proposed technique.

The KNN uses Euclidean distance as the similarity measure between the online query and the FP database. Let  $S$  be the online query from the test point, and  $\delta_i$  is the RSS data from the FP database  $\lambda$ .

$$\xi(i) = \sqrt{(S - \delta_i)^2} \quad (25)$$

$$P_{Est} = \frac{\sum_{i=1}^K (x_i, y_i)}{K} \quad (26)$$

In the probabilistic method, the RSS values obtained are treated as a random variable. For simplicity we represent the RPs here by  $X = (x_1, x_2, \dots, x_N)$ . The working of the probabilistic technique can be stated as we have to find the location in the radiomap that maximizes the conditional probability  $P(x_i|s)$ .

$$P(x | s) = \frac{(P(s | x)P(x))}{P(s)} \quad (27)$$

Without prior information about the position, we can assume that the probability of wireless device located at different places are equally likely.

$$P(x | s) = c \cdot P(s | x) \quad \text{where} \quad c = \frac{P(x)}{P(s)} \tag{28}$$

since  $P(s)$  is constant for all  $x$  and 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  $P(x)$  can be factored out from the maximization process, and the equation becomes

$$\text{argmax}_x [P(x | s)] = \text{argmax}_x [P(s | x)] \tag{29}$$

The location with the maximum probability of occurrence of the RSS vector will be the estimated location. The remaining term  $P(s | x)$  is calculated using the radiomap, where the RSS values are stored as a joint probability distribution.

### 4.1 Positioning Comparison of Algorithm A and Algorithm B

AP1 is placed at  $[50, 1]$ ,  $CK_B$  is kept 2,  $T_{APs}$  is 6, and grid size equal to 3 m and 2 m. After applying both algorithms A and B We analyze the positioning performance using cumulative distribution function. Figure 6 shows the performance of both the algorithms with 3 m grid and KNN, where as in Fig. 7, the grid size is changed to 2 m. Probabilistic matching technique is also used for performance analysis and the results can be seen in Fig. 8. Figure 9 shows the results when the coordinates of AP1 are changed to  $[25, 4]$ . The detailed results of these graphs can be seen in Table 1. From these results; we observe enhanced algorithm performs better than our previous positioning algorithm.

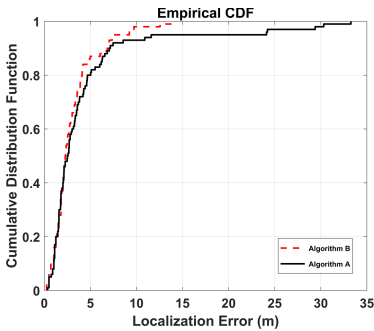


Fig. 6. Deterministic algorithm used with grid size of 3 \* 3.

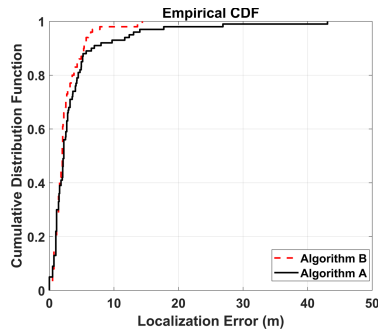
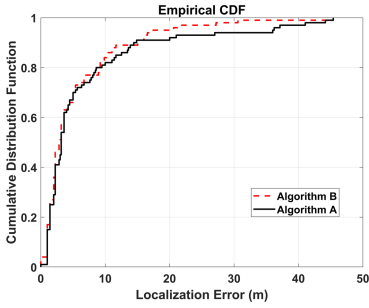
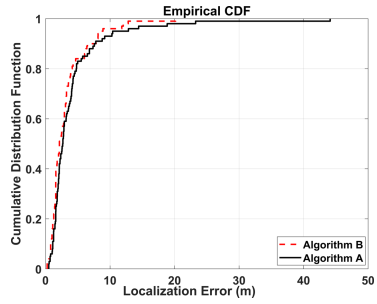


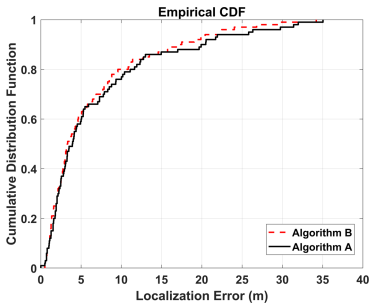
Fig. 7. Deterministic algorithm used with grid size of 2 \* 2.



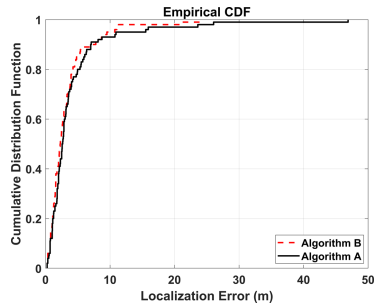
**Fig. 8.** Probabilistic algorithm used with grid size of  $3 * 3$ .



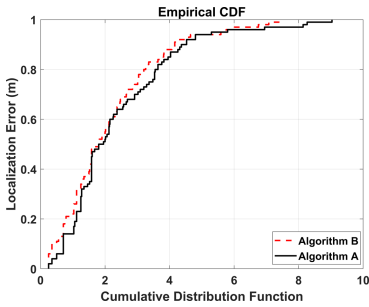
**Fig. 9.** Deterministic algorithm used with grid size of  $3 * 3$ ,  $AP_1 = [25, 4]$ .



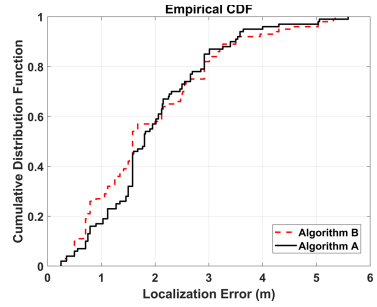
**Fig. 10.**  $CK_B = 1, AP_s = 4$



**Fig. 11.**  $CK_B = 2, AP_s = 6$



**Fig. 12.**  $CK_B = 3, AP_s = 10$



**Fig. 13.**  $CK_B = 4, AP_s = 12$

### 4.2 Impact of Variable $CK_B$ on the Results

The impact of  $CK_B$  on the density of APs and positioning performance can be seen in Fig. 10, 11, 12 and 13 and the details in Table 2. It is a trade-off between positioning accuracy and cost of the system, which can be adjusted according to the requirement and budget. Increasing the value of  $CK_B$  ensures

**Table 1.** Positioning performance (Fig. 7, 8, 9 and 10)

| Figure | Algorithm | Mean error (m) | Maximum error (m) |
|--------|-----------|----------------|-------------------|
| 7      | A         | 4.2718         | 33.2650           |
| 7      | B         | 3.0580         | 13.9754           |
| 8      | A         | 3.6564         | 43.0726           |
| 8      | B         | 2.5426         | 14.3962           |
| 9      | A         | 6.8562         | 45.3982           |
| 9      | B         | 5.5804         | 45.3982           |
| 10     | A         | 4.1065         | 44.1163           |
| 10     | B         | 3.1996         | 20.1556           |

**Table 2.** Impact of variable  $CK_B$ 

| $CK_B$ | No. of APs | Mean error (m) | Maximum error (m) |
|--------|------------|----------------|-------------------|
| 1      | 4          | 6.4            | 34.21             |
| 2      | 6          | 3.3104         | 24.1350           |
| 3      | 10         | 2.1            | 9                 |
| 4      | 12         | 1.93           | 5.3               |

better coverage as a larger number of APs cover a single RP, but too many APs covering every point will also lead to more interference.

The optimization algorithm deploys the APs according to the environmental dynamics and keeping in view the crowd presence. As the crowd profoundly influence the overall optimization process, which is considered in our optimization process as people attenuation factor.

## 5 Conclusion

In this paper, we proposed an AP optimization scheme for indoor environments within the framework of fingerprinting localization systems. This optimization improves the positioning accuracy which can be used for both LBS and CR. The scheme aims to maximize coverage and localization accuracy. A variable is initialized to decide the optimal density of APs, depending upon the requirement. The proposed optimization scheme uses the idea of clusters, based on overlapping regions of APs to select the optimal configuration. This cluster-based technique calculates the objective function cluster-wise. We also analyze the impact of the variable used for optimal AP density on localization accuracy and APs density. To validate the performance of our enhanced optimization algorithm, deterministic and probabilistic localization algorithms are used. The simulation results prove the effectiveness of our newly proposed optimization scheme.

**Acknowledgment.** This work was supported by the National Natural Science Foundation of China under Grants 61771163, the Natural Science Foundation for Outstanding Young Scholars of Heilongjiang Province under Grant YQ2020F001, the Science and Technology on Communication Networks Laboratory under Grants SXX19641X072 and SXX18641X028.

## References

1. Khelifi, F., Bradai, A., Benslimane, A., Rawat, P., Atri, M.: A survey of localization systems in internet of things. *Mob. Netw. Appl.* **24**(3), 761–785 (2018). <https://doi.org/10.1007/s11036-018-1090-3>
2. Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M.: Internet of things for smart cities. *IEEE Internet Things J.* **1**(1), 22–32 (2014). <https://doi.org/10.1109/JIOT.2014.2306328>
3. Khan, A.A., Rehmani, M.H., Rachedi, A.: When cognitive radio meets the internet of things? In: 2016 International Wireless Communications and Mobile Computing Conference (IWCMC), Paphos, pp. 469–474 (2016). <https://doi.org/10.1109/IWCMC.2016.7577103>
4. Li, F., Lam, K., Li, X., Sheng, Z., Hua, J., Wang, L.: Advances and emerging challenges in cognitive internet-of-things. *IEEE Trans. Industr. Inf.* **16**(8), 5489–5496 (2020). <https://doi.org/10.1109/TII.2019.2953246>
5. Höyhtyä, M., et al.: Database-assisted spectrum sharing in satellite communications: a survey. *IEEE Access* **5**, 25322–25341 (2017). <https://doi.org/10.1109/ACCESS.2017.2771300>
6. Xue, W., Qiu, W., Hua, X., Yu, K.: Improved Wi-Fi RSSI measurement for indoor localization. *IEEE Sens. J.* **17**(7), 2224–2230 (2017). <https://doi.org/10.1109/JSEN.2017.2660522>
7. Yoo, J.: Change detection of RSSI fingerprint pattern for indoor positioning system. *IEEE Sens. J.* **20**(5), 2608–2615 (2020). <https://doi.org/10.1109/JSEN.2019.2951712>
8. Raschellà, A., Bouhaf, F., Seyedehbrahimi, M., Mackay, M., Shi, Q.: Quality of service oriented access point selection framework for large Wi-Fi networks. *IEEE Trans. Netw. Serv. Manag.* **14**(2), 441–455 (2017). <https://doi.org/10.1109/TNSM.2017.2678021>
9. Jaiyeola, M.O., Young, M., Xiao, J., Medal, H., Grimes, G., Schweitzer, D.: Towards scalable planning of wireless networks. In: 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Arlington, VA, USA, pp. 629–633 (2019)
10. Raschellà, A., et al.: A dynamic access point allocation algorithm for dense wireless LANs using potential game. *Comput. Netw.* **167**, 106991 (2020)
11. Liu, P., Meng, X., Wu, J., Yao, M., Tang, Z.: AP deployment optimization for WLAN: a fruit fly optimization approach. In: 2019 IEEE/CIC International Conference on Communications in China (ICCC), Changchun, China, pp. 478–483 (2019). <https://doi.org/10.1109/ICCCChina.2019.8855912>
12. Jia, M., Khattak, S.B.A., Guo, Q., Gu, X., Lin, Y.: Access point optimization for reliable indoor localization systems. *IEEE Trans. Reliab.* **69**(4), 1424–1436 (2020). <https://doi.org/10.1109/TR.2019.2955748>