An Automatic Calibration of Large-Scale Indoor Target Signal Positioning System

Zhitao Zhang, Ye Yin, Haoqian Wei, and Wei Tang
College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, 200234, China
Email: 1000359922@smail.shnu.edu.cn; yinye@188.com; {dicktl, weihaoqian492}@sina.cn

Abstract —Traditional indoor localization method based on the received signal strength (*RSSI*) has many drawbacks, such as the low localization accuracy, the high hardware requirements, and the complex algorithm. It is essential to create a robust system that can be suitable for varying indoor environment. In order to solve this problem, an automatic calibration indoor target signal positioning system is proposed in the paper, it can be used to process large-scale and real-time positioning. The system is based on the analysis of large amount measurement during the intensive experiments, mainly including low-complexity Gaussian Filter (GF), Virtual Calibration Technology (VCT), Probabilistic Positioning Algorithm (PPA) and People Pace Model (PPM). In the actual indoor environment, the algorithm is verified by MATLAB simulation. The results show that 100% of the average position error is 57cm, and 70% of the average position error is 15cm, this indoor localization method can be applied to the actual location environment.

Index Terms—Indoor positioning, *RSSI*, automatic calibration, large-scale localization, people pace model

I. Introduction

Since the concept of positioning navigation was proposed, the global positioning system (GPS) has been the earliest known in [1]. Today, with the rapid development of mobile Internet, smart city, and the widespread popularity of smart mobile devices, positioning navigation has produced a huge economic benefit and widely used in scientific research. In [2], indoor positioning system based on the Radio Frequency Identification (RFID) is one of the most typical representatives.

Traditional researches in RFID indoor positioning system are focus on the accuracy of positioning and real-time performance, none of them consider the adaptive capacity of the system in [3]. In [4], most of traditional indoor positioning system is based on LANDMARC, while the positioning accuracy is limited to 1 m to 2 m. In the same time, there are many common positioning algorithm, such as DV-Hop positioning algorithm in [5], k-Nearest neighbor (k-NN) positioning algorithm in [6], Support Vector Machine (SVM) algorithm in [7] and a high degree of complexity of neural network algorithm in [8], they are mainly applied for small and medium indoor environment. However, in the actual demand, large and medium scale indoor positioning environment is desired.

In this paper, we propose an automatic calibration procedure of the signal propagation model that is only based the *RSSI*s measured among readers and that can be executed periodically and automatically. Low-complexity GF is used to process the *RSSI*s before calibration procedure, and the Probabilistic Positioning Algorithm (PPA) based on *RSSI*s works for estimate the locations of objects. Considering the uncertainties caused by the varying environment, the automatic virtual calibration and Bayesian inference is proposed to improve the localization accuracy. This RF-based indoor positioning system is easy to deploy and cost-effective, it can periodically and automatically calibrate the propagation model parameters without human intervention and additional equipment.

II. System Structure and Experimental Environment

In our positioning system, the off-the-shelf long distance active RFID system is used. The system works at the range of 2.4GHz frequency, with a minimal range of 0.5 meter and maximum range of 80 meters. The reader can not only receive *RSSI* from every tag within its range, but also broadcast *RSSI*. Each reader can detect up to 200 tags simultaneously, each tag is pre-programmed with a unique 9-character ID (Identity) for identification by readers. The distance between the readers and tags or readers and readers can be estimated by the propagation model parameters using the RSSI change.

A. Hardware and Network Structure

The system consists of three network layers—RFID sensor layer, data collection layer and data processing layer. The RFID sensor layer is used to transmit power signals to the readers (each target and reference object wears an active tags). The data collection layer is used RFID readers to collect power signals transmitted by

Manuscript received October 30, 2015; revised June 16, 2016.
This work was supported by the Foundation Name under Grant No. 61503251, the Alliance Project of Shanghai No. LM201352, and the School-enterprise Cooperation Project of Qingpu No.2015-19
Corresponding author email: ansandy@foxmail.com
doi:10.12720/jcm.11.6.598-608

RFID sensors, and all readers also can collect their power signals transmitted by each other, this collected data are transmitted to `data processing layer. The data processing layer receives *RSSI* and processes the data for positioning. In practice, the whole detection area may be covered by many tags, readers and servers. For simplicity, Fig. 1 is the hierarchical architecture.

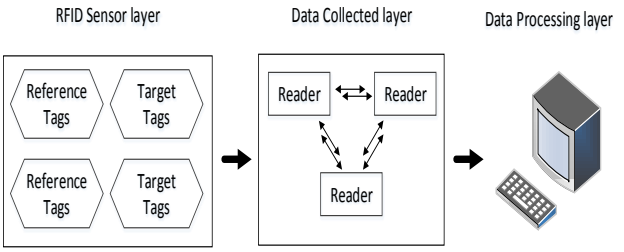


Fig. 1. System structure

In this indoor positioning system mainly includes three kinds of devices: RFID tags, RFID readers and servers. The active tags is shown in Fig. 2, the Reader board is shown in Fig. 3. Table I is the Manufacturer and parameters of them

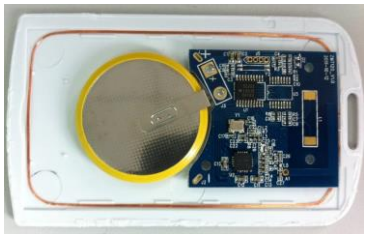


Fig. 2. Tag board

.Fig. 3. Reader board.

TABLE I: MANUFACTURER AND PARAMETERS OF READER AND TAG

	Reader	Tag
Working frequency	2.4GHz	
modulation mode	MSK	
communication distance	0.5meter~80meter	
communication rate	250kbps	
Working voltage	12V	1.5V
Communication interface	RJ45、RS232、RS485	
working temperature	-40℃~+80℃	
Manufacturer: Shanghai Zhen Zhuo Electrical Technology Ltd. Co		

B. Experimental Environment

The experimental environment is in the actual laboratory environment laboratory, the size of the laboratory is about 20m length, 8m width, 3.5m height. In the laboratory, it lay some office segment baffle lattice in per 1.5 square meters, and place some 2.5 meters high experimental equipment cabinet, several experimental platforms. The target in laboratory environment moved in random, some reference tags are also deployed. The laboratory environment is shown in Fig. 4.

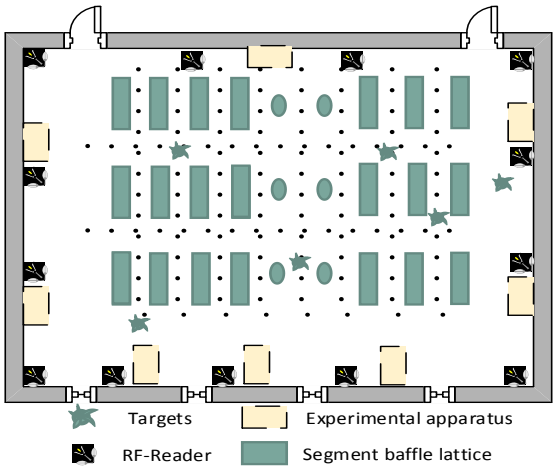


Fig. 4. Experimental environment

III. PRELIMINARY EXPERIMENTS ANALYSIS AND PROCESSING

In this section, the preliminary experiments analysis and processing are the main work. The data collection layer has been discussed, mainly including the positioning of the tag cards and readers, readers and readers. However, the *RSSI* can be affected by the complex indoor environment. So the rough data should be processed firstly. In this paper, a kind of low complexity Gauss filter is proposed to process data.

In fact, the relationship between the signal strength and the distance is not very clear, and there is a lot of fluctuation of *RSSI* in the environment. During a certain period of time, the RFID readers can read a series of *RSSI* values. However, the non-line-of-sight and multipath effects cause these *RSSI* values with a lot of fluctuation. In [9], filtering of *RSSI* is necessary to obtain a group of highly credible *RSSI* values.

In the data collection layer, due to the fluctuation of *RSSI*, the data usually should be collected repeatedly. In [10], L. N. Chen suppose that each measurement is independent, and assume that the distribution of *RSSI* values is a normal distribution. So the Gaussian filter is a good way to process The values of *RSSI* is smooth after the Gaussian filter processing, it improved the positioning accuracy, and the *RSSI*s' distribution is a normal distribution, its mean is μ, and standard deviation is σ. For a specific set of *RSSI* values, expressed as $\{RSSI^t, t=1,2,3...n\}$, which is received at a fixed position from time slot 1 to slot n. Assume that $x \in RSSI^t$, $f(x)$ is to represent Probability Density Function (PDF) of x in [11], given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (1)$$

where n is the times of repeat measurements, $\mu = \frac{1}{n}\sum_{t=1}^{n} RSSI^t$, $\sigma = \sqrt{\frac{1}{n\text{-}1}\sum_{t=1}^{n}(RSSI^t - \mu)^2}$.

In [12], M. Zhu presented the high probability zone as $P>0.6$, which is an experiment value, it is also used in this paper. Through the above analysis, the interval of *RSSI* is [0.26σ+μ, 3.09σ+μ]. The *RSSI* data can be define as a matrix R_e, it is a model matrix that contains 200 sample values, and R_{2m} and R_{3m} are real data matrices that each contain 200 actual *RSSI* values (values are collected at the distances from tag to reader are 2 meters and 3 meters).

$$R_e = \begin{bmatrix} RSSI^1 & RSSI^2 & \dots & RSSI^9 & RSSI^{10} \\ RSSI^{11} & RSSI^{12} & \dots & RSSI^{19} & RSSI^{20} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ RSSI^{191} & RSSI^{192} & \dots & RSSI^{199} & RSSI^{200} \end{bmatrix}_{20\times10}$$

$$R_{2m} = \begin{bmatrix} -67 & -67 & -67 & -67 & -66 & -67 & -68 & -67 & -67 & -68 \\ -67 & -67 & -67 & -67 & -67 & -67 & -67 & -68 & -67 & -67 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -66 & -66 & -67 & -66 & -67 & -66 & -65 & -65 & -66 & -65 \\ -65 & -67 & -67 & -67 & -67 & -64 & -65 & -66 & -67 & -67 \end{bmatrix}_{20\times10}$$

$$R_{3m} = \begin{bmatrix} -72 & -71 & -73 & -77 & -75 & -73 & -75 & -74 & -74 & -74 \\ -71 & -74 & -72 & -71 & -71 & -72 & -70 & -71 & -71 & -72 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -74 & -74 & -74 & -73 & -74 & -71 & -74 & -71 & -74 & -72 \\ -73 & -71 & -71 & -74 & -72 & -75 & -74 & -73 & -71 & -74 \end{bmatrix}_{20\times10}$$

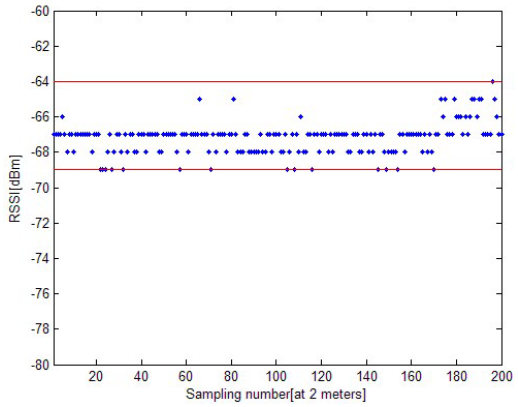


Fig. 5. RSSI Sampled values (two meters)

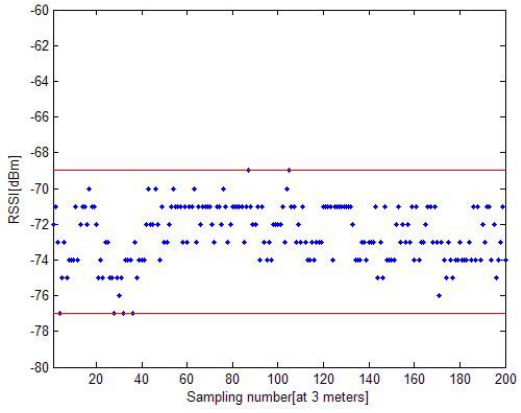


Fig. 6. *RSSI* Sampled values (three meters)

Fig. 5 and Fig. 6 denote two sets of *RSSI* values for which the distances from tag to reader are 2 meters and 3 meters, respectively. The number n of *RSSI* values in each set is 200. From Fig. 5 and Fig. 6, it visually see that fluctuation of samples at 3 meters is larger than samples at 2 meters, the maximum fluctuation amplitude of Fig. 5 is 5 *dBm*, and the maximum fluctuation amplitude of Fig. 6 is 8 *dBm*. This confirms our conclusion in the previous section.

As we know, above two set of values are unfiltered data (or called original data). Our filtering process is to eliminate those values which do not meet the interval [0.26σ+μ, 3.09σ+μ]. This interval is to use unfiltered data calculated out. After filtering, the rest of data is filtered data.

IV. Automatic Calibration Procedure

A. Global Automatic Calibration Procedure

In [13]-[15], most researchers model the indoor path loss with the one-slope model, which assumes a linear dependence between the path loss (*dBm*) and the logarithm of the distance *d* between the object (tag) and reader:

$$PL(d)|_{dBm} = PL(d_0) + 10\eta\log(d/d_0) + WAF \quad (2)$$

In [16] proposed that WAF can be expressed as:

$$WAF = \sum_{i=1}^{N} k_i l_i \quad (3)$$

where k_i is the number of penetrated wall of type *i*, and l_i is the attenuation of the wall of type *i*. The received power *RSSI* is obtained as the difference between the transmitted power P_t and $PL(d)$, *i.e.*

$$RSSI = P_t - PL(d)|_{dBm} \quad (4)$$

substituting eq. (2) into (4) yields (5),

$$RSSI = P_t - PL(d_0) - 10\eta\log(d/d_0) - \mathrm{WAF} \quad (5)$$

Letting $\alpha_0 = P_t - PL(d_0)$ be the *RSSI* at the reference distance of 1 meter. Equation (5) can be expressed as:

$$RSSI = \alpha_0 - 10\eta\log(d/d_0) - \sum_{i=1}^{N} k_i l_i \quad (6)$$

Eq. (6) will be used during the calibration procedure to estimate the propagation model parameters (α_0, η, l_i). Our calibration procedure is independent of the propagation model. In fact, it show that this method are able to estimate the parameters of the chosen propagation model by exploiting the communications among readers, without performing a preliminary measurement phase.

During the virtual calibration procedure, we estimated all the required parameters (α_0, η, l_i) using the Eq. (6). The Global Virtual Calibration Procedure (G-VCP) considers that l_i of all types of walls are the same (l_w). Therefore, in Eq. (6), replacing d (i, j) with the actual distance between $reader_i$ and $reader_j$, k (i, j) with the number of walls crossed by the direct path between $reader_i$ and $reader_j$, we get a estimation $RSSI'(i, j)$:

$$RSSI'_{(i,j)} = \alpha_0 - 10\eta \log(d_{(i,j)}) - k_{(i,j)} l_w \quad (7)$$

$$\forall i, j : (reader_i, reader_j) \in C$$

where, it define C = {$reader_i$, $reader_j$}, $reader_i$ and $reader_j$ represent any two communicated readers. The estimated *RSSI'* (*i, j*) differs from the actual measured *RSSI* (*i, j*) by an error item ε (*i, j*), we assume that all ε (*i, j*) are independent and identically distributed. Keep in mind that α_0 is a priori, according to the approximation of the remaining two parameters (η, l_w) can be achieved by a direct method that minimized the minimum mean square error (MMSE):

$$\varepsilon(i,j) = \left\| RSSI'_{(i,j)} - RSSI_{(i,j)} \right\|_2 \quad (8)$$

the computational overhead for direct method processing a linear MMS estimator problem is polynomial.

B. Calibration Performance Analysis

Let us assume that there are q types of walls in the positioning environment, define $L = \{l_1, l_2, \ldots, l_q\}$, L is the set of attenuation factor for each wall. So, Equation (6) becomes:

$$RSSI'_{(i,j)} = \alpha_0 - 10\eta \log(d_{(i,j)}) - \sum_{d=1}^{q} k_{d(i,j)} l_d \quad (9)$$

$$\forall i, j : (reader_i, reader_j) \in C$$

where $k_{d\,(i,\,j)}$ is the number of wall of type l_d crossed by the direct path between $reader_i$ and $reader_j$ in Eq. (9). The η (path loss exponent) used in this procedure is previously estimated by the G-VCP. Therefore, in our Per-wall Virtual Calibration Procedure (P-VCP), we only estimated the parameters $l_i \in L$, the approximation of the remaining q parameters (l_1, l_2, l_3,..., l_q) also can be achieved by the same method.

In order to evaluate the performance of the calibration methods, in this section, we discuss the estimated distance error after VCT (G-VCP and P-VCP). To facilitate the analysis, we introduce the estimated distance of ideal propagation model for comparison, which is called realistic physical calibration procedure (R-PCP), details refer to. The followings are the estimated distances of these three procedures. Each procedure has two equations, the first equation is the *RSSI* formula, and the second equation is the estimated distance.

(1). R-PCP

$$RSSI = \alpha_0 - 10\eta \log(d) - N(0, \sigma^2) \quad (10)$$

$$d_{R-procedure} = 10^{\frac{\alpha_0 - RSSI - N(0,\sigma^2)}{10\eta}} \quad (11)$$

where d is the distance between reader and tag, $N(0, \sigma^2)$ is the interference term of path loss model in, α_0 is the *RSSI* at distance of 1 meter, η is the path loss exponent. $d_{R\text{-}procedure}$ is the estimated distance between $reader_i$ and tag_j.

(2). G-VCP

$$RSSI = \alpha_0 - 10\eta \log(d) - k_{(i,j)} l_w \quad (12)$$

$$d_{G-procedure} = 10^{\frac{\alpha_0 - RSSI - k_{(i,j)} l_w}{10\eta}} \quad (13)$$

where d is the distance between reader and tag, $k_{(i,\,j)}\, l_w$ is the interference term of path loss model, α_0 is the *RSSI* at distance of 1 meter, η is the path loss exponent. Keep in mind that $k_{(i,\,j)}$ is the number of walls crossed by the direct path between $reader_i$ and tag_j. $d_{G\text{-}procedure}$ is the estimated distance between $reader_i$ and tag_j.

(3). P-VCP

$$RSSI = \alpha_0 - 10\eta \log(d) - \sum_{d=1}^{q} k_{d(i,j)} l_d \quad (14)$$

$$d_{P-procedure} = 10^{\frac{\alpha_0 - RSSI - \sum_{d=1}^{q} k_{d(i,j)} l_d}{10\eta}} \quad (15)$$

where d is the distance between reader and tag, and $\sum_{d=1}^{q} k_{d(i,j)} l_d$ is the interference term of path loss model, α_0 is the *RSSI* at distance of 1 meter, η is the path loss exponent. Keep in mind that $k_{(i,\,j)}$ is the number of walls crossed by the direct path between $reader_i$ and tag_j. $d_{P\text{-}procedure}$ is the estimated distance between $reader_i$ and tag_j.

After the analysis presented above, we analyze the estimated distance. The main purpose of our error analysis is to obtain a distribution of difference value, and select a method with a smallest error. We define following formulas to describe the distance errors:

$$d_{error-R} = d_{practical} - d_{R-procedure} \quad (16)$$

$$d_{error-G} = d_{practical} - d_{G-procedure} \quad (17)$$

$$d_{error-P} = d_{practical} - d_{P-procedure} \quad (18)$$

where $d_{error\text{-}R}$, $d_{error\text{-}G}$, $d_{error\text{-}P}$ are estimated distance errors, $d_{practical}$ is the actual distance between reader and tag.

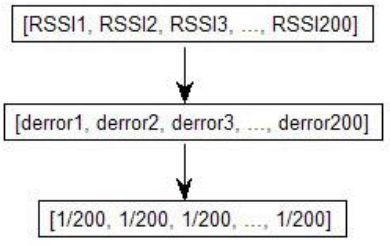


Fig. 7. Flow chart of error analysis

Since each *RSSI* value corresponds to an estimate of the distance, it also corresponds to an estimation error. We assume that each set of *RSSI* values is a uniformly distribution as well as estimation errors. In this part, we just analyze a single value based on this point of view, so as to distinguish it. Fig. 7 is the flow chart of error analysis.

Based on the set of all the *RSSI*s measured between reader and tag, this set of *RSSI*s is R_{2m}. We evaluated the CPD of the estimated distance errors.

Fig. 8 shows that there are three cumulative probability function curves, each curve representing a calibration method. This figure highlights that the P-VCP performs better than the G-VCP and R-PCP. Unsurprisingly, the P-VCP outperforms the G-VCP, due to the better accuracy in the wall modeling. We can also know that virtual calibration procedure results in small estimated distance errors because of more expensive and complicated R-PCP. Table II shows intuitive results of calibration performance.

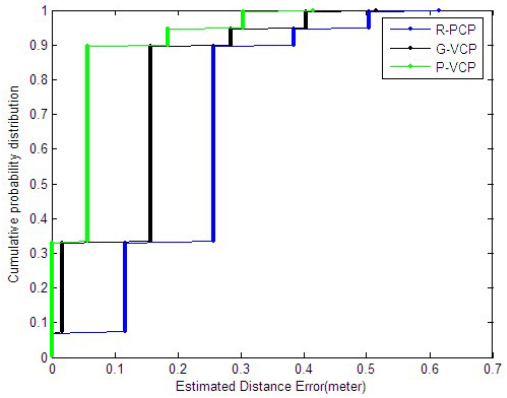


Fig. 8. CPD of estimated distance errors

TABLE II: CPD OF ESTIMATED DISTANCE ERRORS

CPD	P-VCP	G-VCP	R-PCP
0.2	0	2cm	12cm
0.4	5cm	15cm	27cm
0.9	20cm	30cm	39cm

V. PROBABILISTIC POSITIONING ALGORITHM

A. Probability Model of Path Loss

In this section, we establish the probability model of propagation path loss formula, and previous section introduces the propagation path loss model of our experiment.

However, there are a lot of interferences from outside in the practical propagation. For example, walls, elevators, furniture, human activities, multipath effects, shadow effects, etc. In [17], Eq. (2) can be converted into:

$$PL(d)\big|_{dBm} = PL(d_0) + 10\eta\log(d/d_0) + N(0,\sigma^2) \quad (19)$$

where N $(0, \sigma^2)$ denotes a Gaussian distribution with mean of 0 and variance of σ. As we know that *RSSI* is the received power, P_t is the transmit power, we have

$$RSSI = P_t - PL(d)\big|_{dBm} \quad (20)$$

$$RSSI = P_t - PL(d_0) - 10\eta\log(d/d_0) - N(0,\sigma^2) \quad (21)$$

As we know, P_t-$PL(d_0)$ is a *priori*, it is α_0. Then, Eq. (24) can be changed as:

$$RSSI = \alpha_0 - 10\eta\log d - N(0,\sigma^2) \quad (22)$$

where *RSSI* could be regarded as a Gaussian distribution with the mean of (α_0-$10\eta logd$) and variance of σ^2. The PDF of *RSSI* is given by:

$$p(RSSI) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(RSSI-\alpha_0+10\eta\log d)^2}{2\sigma^2}} \quad (23)$$

Through Eq. (2), we have

$$d_{estimate} = 10^{\frac{\alpha_0 - RSSI}{10\eta}} \quad (24)$$

Then Eq. (26) becomes Eq. (28)

$$p(d_{estimate}) = \frac{10\eta}{ln10 d_{estimate}\sigma\sqrt{2\pi}} e^{\frac{-(10\eta\log\frac{d_{actual}}{d_{estimate}})^2}{2\sigma^2}} \quad (25)$$

The Fig. 9 depicts the PDF of the estimated distance. From Fig. 9, we can intuitively draw the following conclusions: 1). If keep variance σ^2 and actual distance d_{actual} unchanged, the curve becomes concentrated and steep with the path loss exponent η increasing. 2). If keep variance σ^2 and path loss exponent η unchanged, the dispersion and smoothness with the actual distance d_{actual} increase. 3). If keep actual distance d_{actual} and path loss exponent η unchanged, the curve shows that dispersion and smoothness with the variance σ^2 increasing.

If we carefully observe the Fig. 8, there are three lines (line1, line2, line3) in the figure, and three lines have the same path loss exponent η, but each line has a different standard deviation σ. The maximum point of each line will drift to the left when standard deviation σ increased. In other words, if the standard deviation σ of each set of *RSSI* is big, the estimated distance will deviate from the true value. Therefore, a stable set of *RSSI* (with small σ) is very important for our system. Fortunately, we can obtain a relatively stable set of *RSSI* with the help of GF process and VCT procedure.

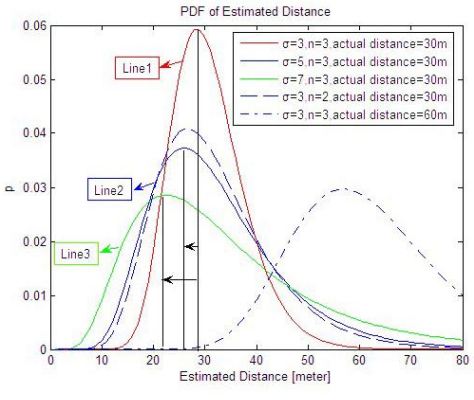


Fig. 9. PDF of the estimated distance

B. People Pace Model

As we know, in our localization procedure, we divide the location area into N grids (see Fig. 9). Each grid has a coordinate corresponding to the center of the grid's coordinate, which is defined as $C_{(gxi,gyi)}$, $i \in (1,2,3...N)$. People move between different grids.

First of all, we study the interval pace of a normal person (interval pace: distance between the adjacent steps, see Fig. 10). In our study, d_1, d_2, d_3,…, d_n are the interval paces. At this stage, we mainly study the distribution of interval paces of different ages.

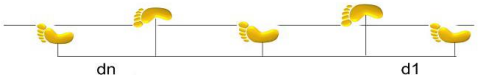


Fig. 10. Pace diagram of one person

Fig. 11 shows four sets of interval paces of different ages. From above figure, we can know that, even for the same person, his (or her) interval paces are different. Looking into the figure above, the maximum fluctuation of each set of interval paces is less than 10cm. Through this experiment, we can draw a conclusion that each people has a different interval pace. In most cases, these values trend to 50cm, and the maximum fluctuation of one person's interval paces is less than 10cm.

We carried out an experiment: there are four people aged 25, 35, 45, 55 years old. Everyone were moving straight up to 100m, and recorded each person's walking step (or called interval pace-IP). Experimental atmosphere is under everyday natural conditions, and there is no interference with emotions, such as anxiety and anger. Statistics are as follows.

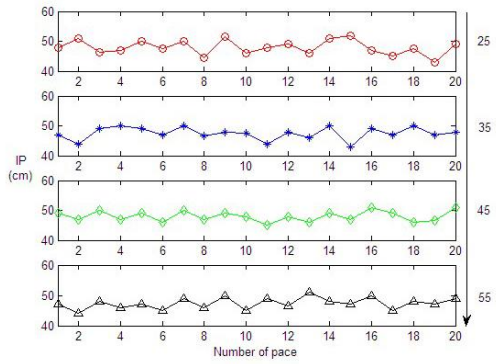


Fig. 11. Pace charts of different people

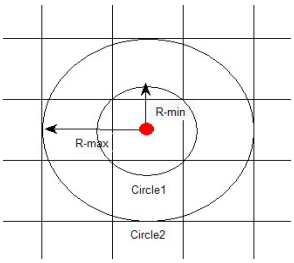


Fig. 12. Span figure of one step

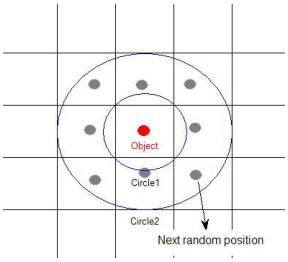


Fig. 13. Next random position

After numerical statistics, we turn to model analysis. Firstly, Fig. 12 shows the location scenario, the solid red circle represents the object, and the location of object is the center of the grid. There are two concentric circles with radius *R-min* and *R-max* which share the common center object. In Fig. 12, *R-min* and *R-max* represent the minimum value (smallest interval pace) and the maximum value (biggest interval pace). Keep this principle: People can go to the adjacent grid at every step (see Fig. 13, the solid gray circle represents the next random position of object after walking a step).

Looking closely at Fig. 12, circle1 is the circumcircle of a grid (each grid is a square), circle2 is the circle inscribed in a square (this square consists of nine grids). In order to determine the size of each grid, we analyzed in two different conditions in [18]:

1. Considering *R-min* is the interval pace

Under this condition, we can achieve the length L_{g1} of grid.

$$L_{g1} = \sqrt{2}R_{min} \tag{26}$$

2. Considering *R-max* is the interval pace

Under this condition, we can also achieve the length L_{g2} of grid.

$$L_{g2} = \frac{2}{3}R_{max} \tag{27}$$

Therefore, we can obtain the range of the length of grid, which is $[0.667R_{max}, 1.414R_{min}]$. According to this interval, we can set up a suitable size of grid, rather than setting up without any basis. This is one of the strengths of our article.

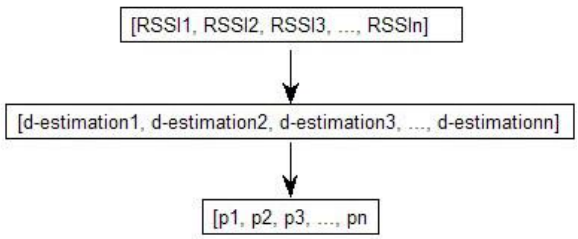


Fig. 14. Flow chart of *RSSI* processing

C. RSSI Processing Method

As we all know, the arithmetic average method is widely used in positioning system for processing the *RSSI*. However, we can know that each *RSSI* value will correspond to a probability in [19]. Therefore, in this section, we propose a weighted average method to process the *RSSI*. Following Fig. 14 is the flow chart of this process.

$[RSSI_1, RSSI_2, RSSI_3, \ldots, RSSI_n]$ is a set of filtered data, $[d_{estimation1}, d_{estimation2}, d_{estimation3}, \ldots, d_{estimationn}]$ is a set of estimated distances corresponding to each *RSSI* value, and $[p_1, p_2, p_3, \ldots, p_n]$ is a set of probabilities corresponded to each estimated distance. The estimated distance can be calculated by Eq. (18), and the probability of each *RSSI* can also be calculated by Eq. (28).

Next, we calculate the weight of a single *RSSI* in the entire set of *RSSI*s.

$$\xi_i = \frac{p_i}{\sum_{i=1}^{n} p_i} \tag{28}$$

Finally, we can obtain the weighted average of *RSSI*s,

$$\overline{RSSI_w} = \sum_{i=1}^{n} \xi_i RSSI_i \tag{29}$$

Table III shows that the weighted average method we proposed is better than traditional arithmetic average method.

TABLE III: COMPARISONS OF TWO DIFFERENT METHODS

Actual instance (meter)	Mean of *RSSI* (dBm)	Estimated distance (meter)	error
Arithmetic average method			
2.0	-68.2239	1.9151	4.24%
3.0	-73.7295	2.9222	2.59%
Weighted average method			
2.0	-68.4221	1.9445	2.78%
3.0	-73.9120	2.9634	1.22%

D. Bayesian Localization Scheme

As we know, the key of tracking and positioning is to get location information. In our system, we define the location information x as a series of coordinate points. In [20], we can obtain the following formula.

$$p\left(x|RSSI^t\right) = \frac{p\left(RSSI^t|x\right)p(x)}{\int p\left(RSSI^t|x\right)p(x)dx} = \frac{p(RSSI^t \mid x)p(x)}{p(RSSI^t)} \tag{30}$$

where $p(x)$ is the priori probability of the object's location. $p(RSSI^t|x)$ can be derived from Eq. (26), which means the PDF of *RSSI* under the situation of location information is known. $p(RSSI^t)$ is the marginal PDF of $p(RSSI^t,x)$. $p(x|RSSI^t)$ is the key equation, we can obtain the location from it. Through eq. (30), we have the following formula.

$$p(x \mid RSSI^t) \propto p(RSSI^t \mid x)p(x)$$

We assume that $Bel(x)$ is the probability of object at location x. $Bel^-(x)$ is the initial probability of object at location x. β is a constant, to normalize the $Bel(x)$ we have

$$Bel(x) = \beta p(RSSI^t \mid x_k)Bel^-(x) \tag{31}$$

$Bel^-(x)$ can be computed from Eq. (35).

$$Bel^-(x) = \int p(x_k \mid x_{k-1})p(x_{k-1} \mid RSSI^t)dx_{k-1} \tag{32}$$

where x_k is the location of object at time k and x_{k-1} is the location of front moment. $p(x_k|x_{k-1})$ is the Dynamic Model of system.

$$p(x_k \mid x_{k-1}) = \begin{cases} 1 & x_k \neq x_{k-1} \\ 0 & x_k = x_{k-1} \end{cases} \tag{33}$$

For a static system, Eq. (35) can be converted into Eq. (34).

$$Bel^-(x_k) = Bel(x_{k-1}) \tag{34}$$

Then we can get our Bayesian estimation model.

$$Bel(x_k) = \beta p(RSSI^t \mid x_k)\mathrm{Bel}(x_{k-1}) \tag{35}$$

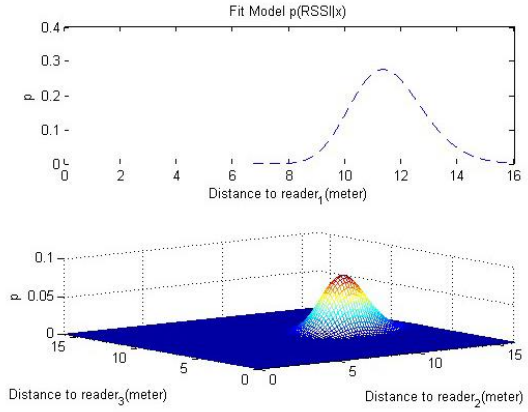


Fig. 15. Fit Model of one tracked tag

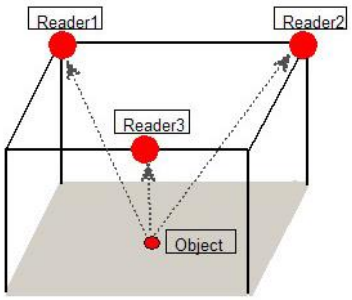


Fig. 16. Deployment environment

We divide the location area into N grids. Looking back to the previous Section, the interval of the length is $[0.667R_{max}, 1.414R_{min}]$, and according to those four sets of interval paces, there is a minimum of 43, and a maximum of 53, then, we can obtain an interval of actual data that is [35.4, 60.8].

In our experiment, we choose 40cm as the length of each grid. The grid size is 40cm×40cm, just like the size of a tile. Each grid has a coordinate corresponding to the center of the grid's coordinate, which is defined as $C_{(gxi,gyi)}$, $i \in (1,2,3...N)$. In this area, the small red dot represents an object tag, and three big solid red circles are readers, That is shown in Fig. 16.

As described above, the initial $p(x)$ could be set as $1/N$, thus the initial value of $Bel(x_k)$ is $1/N$. The $p(RSSI^t|x_k)$ is a Fit Model, which can be calculated by the following formula. Fig. 15 shows the Fit Model of a tracked point.

$$p(RSSI^t \mid x_k) = \prod_{j=1}^{3} p(RSSI_j^t \mid x_k) \tag{36}$$

There are two pictures in the Fig. 15, the precondition of this model is that *RSSI*s are known and unchanged. The upper one represents the probability of the distance from tag to one reader. The lower one represents the distances from tag to the other two readers. From Fig. 16, we can know that was the Fit Model relatively large

values are achieved, when the estimated locations are close to the actual location. And the maximum value is obtained when the estimated location is actual location.

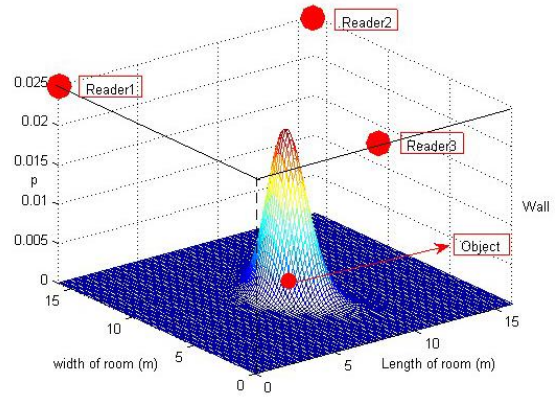


Fig. 17. Fit Model of our environment

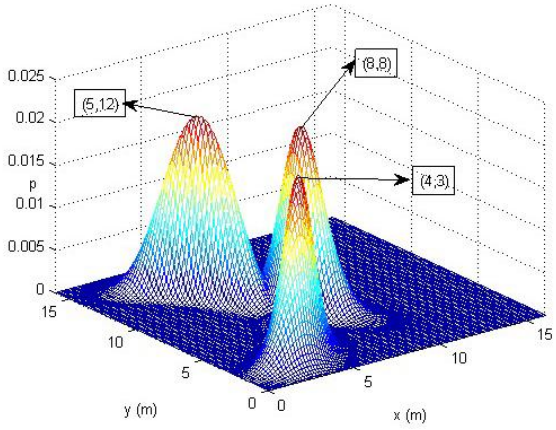


Fig. 18. Fit Model of three tracked tags

Fig. 17 shows a Fit Model of our environment (see Fig. 16). The scene of our experiment is a room with size of 16m×16m, the length of room is marked as X-axis, and the width of room is marked as Y-axis. The coordinate of reader1 is (0, 16, 3), the coordinate of reader2 is (16, 16, 3), the coordinate of reader3 is (8, 0, 3). The coordinate of object tag is (8, 8, 1), this tag is placed on a 1-meter-high triangle bracket. From Fig. 17, we can draw the conclusion that Fit Model can clearly reflect the result of the positioning procedure.

We tracked three objects in this room (see Fig. 18), and the plane coordinates of these three objects have been marked out in the figure. From Fig. 18, the result of positioning procedure is still very clear, even there are multiple targets in the same room.

At the end of algorithm, we evaluate the credibility of output by defining a confidence function $Bel'(x_k)$.

$$Bel'(x_k) = \frac{Bel(x_k)}{max(Bel(x_k = (x_k, y_k)))} \quad (37)$$

We set a threshold T ($0 \le T \le 1$) and only choose those grids with $Bel'(x_k) > T$. After one recursion, an estimated area is obtained.

Fig. 19 shows sectional drawing of space corresponding to Fig. 18. There are three objects in environment. The white areas are estimated areas of tracked objects. There are four pieces of sectional drawings with different thresholds, and from above figure, we can observe that estimated areas of three tracked objects are smaller as T increasing.

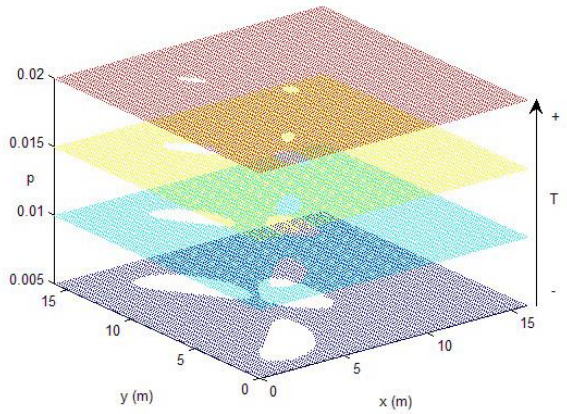


Fig. 19. Sectional drawing of space

Therefore, we can achieve one more precise estimated area, if increasing the recursion time R and threshold T.

VI. System Performance Analysis

A. Average Positioning Distance Error

In this part, we define the *APDE* (Average Positioning Distance Error) to evaluate our system's positioning error, which is given by

$$APDE = \frac{\sum D(C_{(gxk,gyk)}) Bel'(x_k = C_{(gxk,gyk)})}{\sum Bel'(x_k = C_{(gxk,gyk)})} \quad (38)$$

The estimated points are in the estimated area, denoted as (gx_k, gy_k), $k=1,2,3\ldots E$. E is the number of grids covered by the estimated area. (x_0, y_0) is the actual coordinate of object tag. The $D(C_{(gxk,gyk)})$ is denoted as the Euclid Distance from estimated point to the object's actual point. We have

$$D\left(C_{(gxk,gyk)}\right) = \sqrt{(gx_k - x_0)^2 + (gy_k - y_0)^2} \quad (39)$$

The *APDE* is the average error distance between the object and grid points in the estimated area. It represents the accuracy of positioning. When the *APDE* is the minimum value of zero, the system has the highest accuracy. The greater the *APDE* value, the worse the position accuracy.

B. Impact of System Parameters

In this part, we will investigate the system performance depending on a number of system and environment parameters and then simulate the experiment. We use MATLAB to simulate the impact of the following parameters on the localization accuracy, which is represented by average positioning error distance (*APDE*): filtered data and unfiltered data, recursion time R, threshold T, path loss exponent η, and window size w, etc.

We study the impact of filtered data and unfiltered data on average positioning error through experiment 1.

Experiment 1: We assume two scenarios: 1) the input *RSSI* values are unfiltered data. 2) the input *RSSI* values are filtered data. The system is based on the following configuration: grid size L=40cm, threshold T=0.5, path loss exponent η=3, standard deviation σ=1.45, α_0= -59.7581, window size w=200.

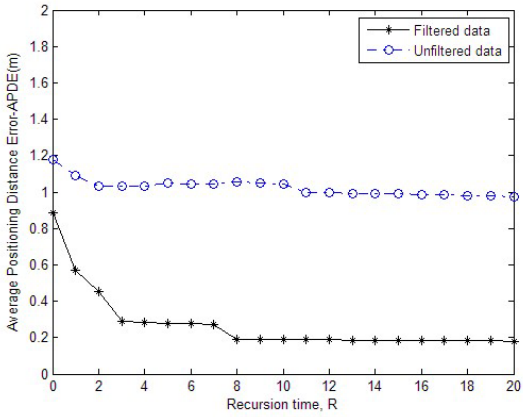


Fig. 20. Impact of Gaussian filter

Fig. 20 depicts *APDE* as a function of the recursion time for the two scenarios. In the first scenario, we observe that the decrease of *APDE*'s rate is very slow as we increase the recursion time. After several rounds of recursive, *APDE* is still large, and the rate of decline in *APDE* is not significant. However, in the second scenario, we observe that *APDE* decreases after increasing the recursion time. In other words, we can draw a conclusion that the use of the Gaussian Filter which will significantly improve the reliability of *RSSI* values.

Fig. 21 illustrates the *APDE* as a function of the recursion time for different objects in the above second scenarios. From Fig. 21, we observe that *APDE* reduced speed is a function of the object locations within the basic detection area. For object 2 and object 3, the localization accuracy increases much faster than that of object 1, with increasing of the recursion time. Generally speaking, the position accuracy is increased with the increasing of the recursion time.

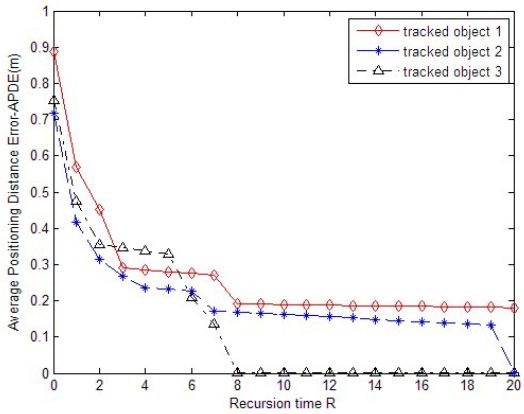


Fig. 21. Tracked three objects

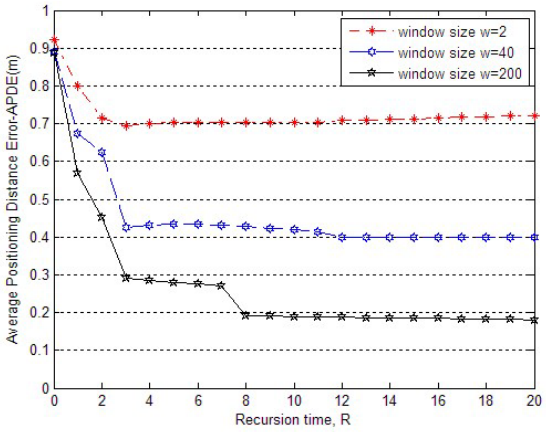


Fig. 22. Impact of window size w

Experiment 2: The following system is configured as: grid size L=40cm, threshold T=0.5, path loss exponent η=3, standard deviation σ=1.45, α_0= -59.7581. Fig. 22 plots *APDE* as a function of the window size *w*. In this part, we plan to evaluate the impact of window size on the localization accuracy. We perform the simulation for object 1, object 2 and object 3 as depicted in Fig. 21.

From Fig. 22, we observe that *APDE* decrease as the window size *w* increase. In another words, the greater the window size *w*, the higher the localization accuracy. We also find that the *APDE* decrease rate depends on the window size. The *APDE* decrease rate will very high when the window size is relatively large.

From Fig. 22, we can also observe a unique phenomenon, when the window size *w* is very small, for example *w* = 2 or 4, the *APDE* decrease in early stage, but it slightly increases in later stage. The main reasons for this phenomenon are sample quantity and fault-tolerant rate. The fault-tolerant rate will increase with a less sample quantity. So, *APDE* will decrease with a bigger sample quantity (or called window size *w*).

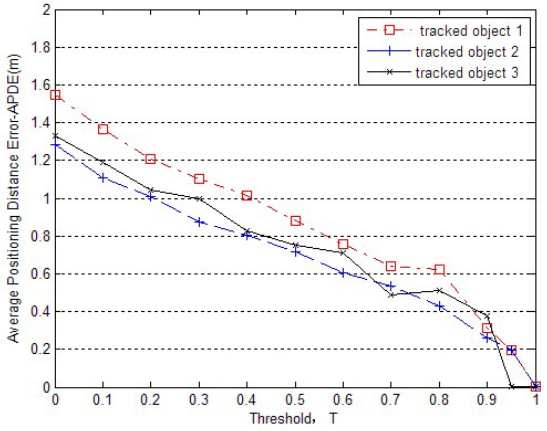


Fig. 23. Impact of threshold T

Experiment 3: The system configured as: grid size L=40cm, path loss exponent η=3, standard deviation σ=1.45, $\alpha 0$= -59.7581. In this experiment, we aim to evaluate the impact of threshold on the localization accuracy. We perform the simulation for object 1, object 2

and object 3. Fig. 24 plots *APDE* as a function of the threshold *T*.

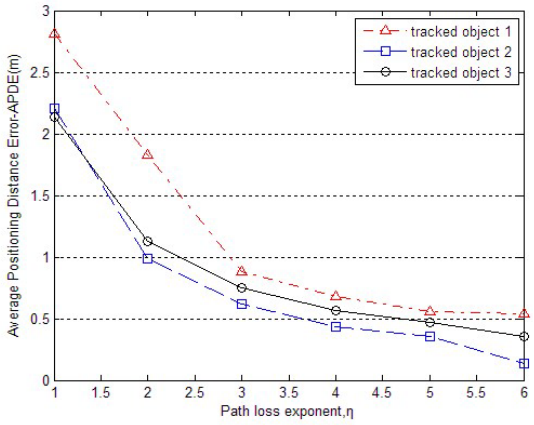


Fig. 24. Impact of path loss exponent η

From Fig. 23, we can get the conclusion that in most situation, the large the threshold *T*, the better the localization accuracy. We also can know that the relationship between threshold *T* and localization accuracy is linearly decreasing. In other words, the localization accuracy can be fully decided by the threshold *T*, and the system will not appear the phenomenon of a steady state of *APDE* (see Fig. 20, Fig. 21, Fig. 22 and Fig. 24, the *APDE* may finally be stable in a value).

Experiment 4: The system configured as: grid size *L*=40cm, threshold *T*=0.5, standard deviation σ=1.45, α_0= -59.7581. In this part, we plan to evaluate the impact of path loss exponent on the localization accuracy. We perform the simulation for object 1, object 2 and object 3. Fig. 24 plots *APDE* as a function of the path loss exponent η.

From Fig. 24, we can see that localization accuracy increase with the increase of path loss exponent η. But in a real environment, the path loss exponent η can't be artificial changed, and in the same environment, η is basically unchanged.

However, from above conclusion, it can give us an instruction that a higher path loss exponent η will reduce *APDE* of our system.

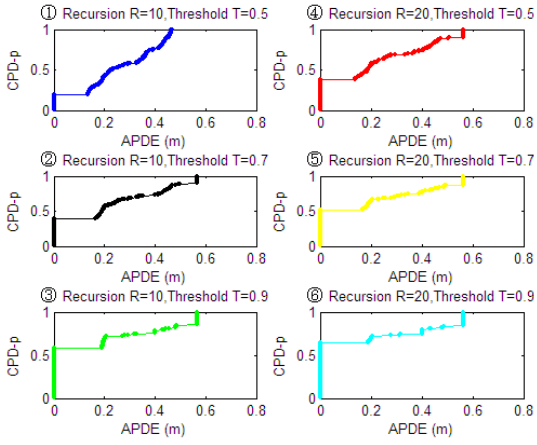


Fig. 25. CPD of different conditions

Experiment 5: This experiment is designed to probabilistically evaluate the system localization performance under the following settings: Using the same reader network topology. We track 169 objects which are uniformly distributed within the basic detection area. The other simulation parameters are: grid size *L*=40cm, path loss exponent η=3, standard deviation σ=1.45, α_0= -59.7581, threshold *T*=0.5, 0.7, 0.9. We test the system performance for two scenarios: 1) recursion time *R*=10; 2) recursion time *R*=20.

Fig. 25 shows the cumulative probability distribution (CPD) obtained from scenarios 1 and 2. There are six figures of different conditions. From the six figures, we can know that in 100% percent of the localization estimation, the system provides objects location with the *APDE* less than 60cm. If we keep threshold *T* unchanged, Fig. 25 shows that the high precision positioning range (in this range, the *APDE* is less than 20cm) increased by 20% with the recursion time increasing from *R*=10 to *R*=20.

VII. Conclusion

In this paper, we presented a new method, easy-setup and cost-effective indoor positioning method based on off-the-shelf active RFID technology. It is used a large number of experimental data to achieve the positioning, and we have evaluated the performance of our proposal by MATLAB simulations. In this system, the simulation results show that the *APDE* are less than 57cm in 100% percent, the *APDE* are less than 35cm in 80%, and the *APDE* are less than 15cm in 70%. The simulation also shows that the system performance improved in the higher values of recursion time, window size, and path loss exponent. The simulation results can prove that the proposed system is an accuracy and cost-effective candidate for future indoor localization.

Acknowledgment

This work is supported in part by National Natural Science Foundation of China under Grant No. 61503251, the Alliance Project of Shanghai No. LM201352, and the School-enterprise Cooperation Project of Qingpu No.2015-19. The authors would like to thank the reviewers for their detailed reviews and constructive comments, which have helped to improve the quality of this paper.

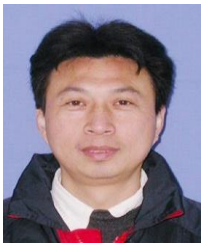
References

[1] D. Dardari, P. Closas, and P. M. Djuric, "Indoor tracking: Theory, methods, and technologies," *IEEE Trans. on Vehicular Technology*, vol. 64, no. 4, pp. 1263-1278, Apr. 2015.

[2] P. Yang and W. Y. Wu, "Efficient particle filter localization algorithm in dense passive RFID tag environment," *IEEE Trans. on Industrial Electronics*, vol. 61, no. 10, pp. 5641-5651, Oct. 2014.

[3] K. M. Yu, M. G. Lee, H. N. Hsieh, *et al.*, "Implementation of an RFID-Based virtual signal mechanism for indoor location sensing system," *Journal of Internet Technology*, vol. 14, no. 4, pp. 631-641, Jul. 2013.
[4] J. H. Li, G. M. Zhang, L. Yu, *et al.*, "An advanced RFID localization algorithm based on region division and error compensation," *Ksii Trans. on* Internet *and Information Systems*, vol. 7, no. 4, pp. 670-691, Apr. 2013.
[5] L. Q. Gui, T. Val, A. Wei, *et al.*, "Improvement of range-free localization technology by a novel DV-Hop protocol in wireless sensor networks," *Ad Hoc Networks*, vol. 24, pp. 55-73, Jan. 2015.
[6] M. A. Bitew, R. S. Hsiao, H. P. Lin, *et al.*, "Hybrid indoor human localization system for addressing the issue of RSS variation in fingerprinting," *International Journal of Distributed Sensor Networks*, 2015.
[7] S. Bai, J. J. Hou, and N. Ohnishi, "Combining LBP and SIFT in sparse coding for categorizing scene images," *Ieice Trans. on Information and Systems*, vol. E97D, no. 9, pp. 2563-2566, Sep. 2014.
[8] S. Kim, J. H. Lee, and J. W. Moon, "Performance evaluation of artificial neural network-based variable control logic for double skin enveloped buildings during the heating season," *Building and Environment*, vol. 82, pp. 328-338, Dec. 2014.
[9] Y. B. Zhao, Y. Yang, and M. Kyas, "Adaptive range-based nonlinear filters for wireless indoor positioning system using dynamic gaussian model," *IEEE Trans. on Vehicular Technology*, vol. 64, no. 9, pp. 4282-4291, Sep. 2015.
[10] L. N. Chen, B. H. Li, K. Zhao, *et al.*, "An improved algorithm to generate a Wi-Fi fingerprint database for indoor positioning," *Sensors*, vol. 13, no. 8, pp. 11085-11096, Aug. 2013.
[11] D. L. Guo, Y. D. Zhang, Q. Xiang, *et al.*, "Improved radio frequency identification indoor localization method via radial basis function neural network," *Mathematical Problems in Engineering*, 2014.
[12] M. Zhu and H. Zhang, "-Research on model of indoor distance measurement based on receiving signal strength," in *Proc. International Conference on Computer Design and Applications*, 2010, pp. 54-58.
[13] S. Tomic, M. Beko, and R. Dinis, "RSS-Based localization in wireless sensor networks using convex relaxation: Noncooperative and cooperative schemes," *IEEE Trans. on Vehicular Technology*, vol. 64, no. 5, pp. 2037-2050, May 2015.
[14] S. Savazzi, M. Nicoli, F. Carminati, *et al.*, "A bayesian approach to device-free localization: Modeling and experimental assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 1, pp. 16-29, Feb. 2014.
[15] J. G. Shang, X. K. Hu, F. Q. Gu, *et al.*, "Improvement schemes for indoor mobile location estimation: A survey," *Mathematical Problems in Engineering*, 2015.
[16] P. Merello, F. J. Garcia-Diego, and M. Zarzo, "Microclimate monitoring of Ariadne's house (Pompeii, Italy) for preventive conservation of fresco paintings," *Chemistry Central Journal*, vol. 6, Nov. 2012.
[17] E. Valero, A. Adan, and C. Cerrada, "Automatic construction of 3D basic-semantic models of inhabited interiors using laser scanners and RFID sensors," *Sensors*, vol. 12, no. 5, pp. 5705-5724, May 2012.
[18] S. Sendra, J. Lloret, C. Turro, *et al.*, "IEEE 802.11a/b/g/n short-scale indoor wireless sensor placement," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 15, no. 1-3, pp. 68-82, 2014.
[19] Y. T. Ma, L. J. Zhou, K. H. Liu, *et al.*, "Iterative phase reconstruction and weighted localization algorithm for indoor RFID-Based localization in NLOS environment," *IEEE Sensors Journal*, vol. 14, no. 2, pp. 597-611, Feb. 2014.
[20] L. Xiao, Y. Yin, X. N. Wu, *et al.*, "A large-scale RF-Based indoor localization system using low-complexity gaussian filter and improved bayesian inference," *Radioengineering*, vol. 22, no. 1, pp. 371-380, Apr. 2013.

Zhitao Zhang was born in Hubei Province, China, in 1989. He received the B.S. degree from Shanghai Normal University (SHNU). He is currently pursuing the M.S. degree in the college of Information, Mechanical and Electrical Engineering, Shanghai Normal University. His research interests include indoor positioning, communication coding.



Ye Yin was born in Shanghai, China, in 1961. He received the B.S., M.S. and Ph.D. degree from Tongji University. Prof. Yin's research interest is currently at information theory, granular computing, fuzzy control, two-dimension barcode and communication coding

Haoqian Wei was born in Shanxi Province, China, in 1992. She received the B.S. degree from Shanghai Normal University (SHNU), Shanghai, in 2014, in communication engineering. She is currently pursuing the M.S. degree with the Communication and Information System, SHNU. Her research interests include Indoor Positioning and two-dimension barcode.

Wei Tang was born in Shandong Province, China, in 1990. He received the B.S. degree from Shanghai Normal University (SHNU). He is currently pursuing the M.S. degree in the college of Information, Mechanical and Electrical Engineering, Shanghai Normal University. His research interests include fuzzy control.