A Large-scale RF-based Indoor Localization System Using Low-complexity Gaussian Filter and Improved Bayesian Inference

Long XIAO[1], Ye YIN[1], Xiangnong WU[1], Jianwei WANG[2]

[1]College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, 200234, China
[2] Shanghai Zhen Zhuo Electrical Technology Ltd. Co., Shanghai, 201100, China

loongshaw@gmail.com, yinye@188.com, xnwu@shnu.edu.cn, wangjianwei@zenkore.com

Abstract. *The growing convergence among mobile computing device and smart sensors boosts the development of ubiquitous computing and smart spaces, where localization is an essential part to realize the big vision. The general localization methods based on GPS and cellular techniques are not suitable for tracking numerous small size and limited power objects in the indoor case. In this paper, we propose and demonstrate a new localization method, this method is an easy-setup and cost-effective indoor localization system based on off-the-shelf active RFID technology. Our system is not only compatible with the future smart spaces and ubiquitous computing systems, but also suitable for large-scale indoor localization. The use of low-complexity Gaussian Filter (GF), Wheel Graph Model (WGM) and Probabilistic Localization Algorithm (PLA) make the proposed algorithm robust and suitable for large-scale indoor positioning from uncertainty, self-adjective to varying indoor environment. Using MATLAB simulation, we study the system performances, especially the dependence on a number of system and environment parameters, and their statistical properties. The simulation results prove that our proposed system is an accurate and cost-effective candidate for indoor localization.*

Keywords

Indoor localization, Gaussian filter, wheel graph, probabilistic localization algorithm.

1. Introduction

The growing of wireless communications, portable computers, and smart sensors has generated much commercial and research interests in statistical methods to track people and things inside stores, hospitals, warehouses, and factories, where Global Positioning System (GPS) [1] devices generally do not work and Indoor Positioning System (IPS) [2], [3], [4] aims to provide location estimation for wireless devices, such as handheld devices and electronic badges. In this paper, we propose and study a new method for indoor localization, which is cost-effective for tracking objects.

Nowadays, due to advantages such as small size, low power, low cost and easy deployment, the Radio Frequency Identification (RFID) [5] sensors are widely used to implement ubiquitous computing and smart space. We have successfully applied this technique to the 2010 Shanghai World Expo [6]. With the capability of providing RSS (Received Signal Strength) information, current advanced RFID systems are one of the potential candidates for indoor localization. Several RFID based systems have been proposed for tracking objects in indoor environments. As we know, SpotON [2], [7] uses an aggregation algorithm for three-dimensional localization. The tags use RSS information to obtain inter-tag distances based on empirical mapping between the two. SpotON assumes deterministic mapping between RSS and distance and does not account for the range measurement uncertainty caused by the varying environment. LANDMARC [2], [8] utilizes RSS measurement information to locate objects using k nearest reference tags. To diminish the uncertainty of the detected range caused by the varying environments, a large number of reference tags must be distributed in the environment. This limits its applications for most indoor scenarios, such as people tracking, fire disaster site, and so on.

In this paper, we present a new method based on off-the-shelf active RFID technology, Gaussian Filter (GF), Wheel Graph Model (WGM) and Probabilistic Localization Algorithm (PLA). It is easy to deploy and cost-effective for indoor localization. Considering the uncertainties caused by the varying environment, we incorporate a probabilistic scheme based on Gaussian filter pretreatment, on-site calibration and Bayesian inference to improve the localization accuracy. Bayesian inference was also used for traditional cellular and WLAN-based localization system. But with an additional WGM, our proposed system can be applied to large-scale indoor localization.

The rest of the paper is organized as follows: Section 2 describes our method, including the basic system architecture, system overview, GF pretreatment, path loss model, WGM and the PLA. In Section 3, the method of performance metric is introduced, followed by simulation results and discussions. And finally in Section 4 we conclude the paper.

2. System Architecture

In our localization system the off-the-shelf long distance active RFID system [9] is used. The system works at the range of 2.4 GHz frequency [10], with a minimal range of 0.5 meter and maximum range of 80 meters. The reader can receive Received Signal Strength Indicator (RSSI) from every tag within its range. Each reader can detect up to 200 tags simultaneously. Each RFID tag is pre-programmed with a unique 9-character ID (Identity) for identification by readers. In the next section, we will discuss the network layout of the system.

2.1 Network Layout

The system consists of three network layers: sensing network layer, data collection layer, processing layer. The sensing network layer is used to measure the RSSI information from the readers to objects (active tags) and to transmit the information to next layer. The data collection layer is used to collect RSSI information and to transmit the data to next layer. The processing layer receives RSSI information and processes the location information. In practice, the whole detection area may be covered by several servers. For simplicity, in Fig. 1, we show the hierarchical architecture within the coverage of one server. Fig. 4 shows the reader and tags used in our experiments.

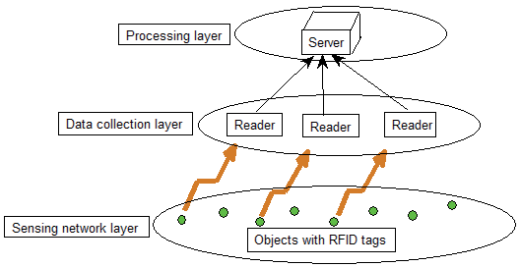

Fig. 1. System structure.

Fig. 2. Reader.

Fig. 3. Tags.

Fig. 4. Reader and tags.

RFID tags: We use the active tags in our system, as shown in Fig. 3. They are deployed in the sensing network layer, and are divided into two categories: object tags and reference tags. Each tracked object will be attached with a unique active RFID tag, called "object tag", used for identifying and tracking objects. Each tag has a unique ID, hence, we can distinguish objects by the corresponding ID number. Reference tags are the active RFID tags used for calibrating environment parameters. The active tags will periodically emit signals with their IDs.

RFID readers: The data collection layer consists of readers. Each reader, as shown in Fig. 2, also has a unique ID number. Every small detection area contains three readers. The whole detection area is covered by the data collection layer. Every object tag should be within the readable range of readers. The principle of readers' deployment should be satisfied that the distribution of the readers is not in one line in the space, and all readers' locations are known. A reader is responsible for: 1) Collecting and decoding the signals emitted by the active tag in its coverage; 2) Measuring the RSSI for each tag within its range; 3) Reporting tag ID, corresponding RSSI, and its own ID number to the server. To realize these functions, each reader is equipped with two interfaces: a RF interface that detects tags within its range, and a communication interface that transmits data to servers.

Servers: Each reader should be within the reach of at least one server. We can see this from Fig. 1. Readers communicate the measured RSSIs of the tags with the server. A server is responsible for: 1) Collection of RSSIs and IDs coming from readers. 2) Calculating the location of the object tags according to the positioning algorithm.

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

Tab. 1. Manufacturer and parameters of reader and tag.

Fig. 2-4 show readers and tags used in the experiment, Tab. 1 gives the Manufacturer and parameters of them.

The principles of the system are as follows. After the system is setup, the RFID readers will begin to detect signals sent by RFID tags. If any tag is located within its coverage, the reader will collect and report the RSSIs of the tag and IDs both of the tag and the reader to the server. All the data reported by the RFID readers will be processed at the server. The server will estimate the locations of the objects according to the localization algorithm using the RSS information and the readers' locations.

2.2 Gaussian Filter Pretreatment

As we know, RSSI is very sensitive to the environment. This will limit the accuracy of the measurement. In fact, the relationship between the signal strength and the distance is not very clear, and there is a lot of volatility 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 volatility. Hence, filtering of RSSI is necessary to obtain a group of highly credible RSSI values.

In the data collection stage, due to the volatility of RSSI, we usually repeat the experiments. Supposing that each experiment is independent, we can assume that the distribution of RSSI values is a normal distribution. Through Gaussian filter [11] processing, we can obtain a set of relatively smooth values of RSSI with improved reliability, and improved positioning accuracy.

Due to the RSSIs' distribution is a normal distribution model, 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 $x \in RSSI^t$, we use $f(x)$ to represent Probability Density Function (PDF) of x, given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \tag{1}$$

where n is the times of repeat measurements,

$$\mu = \frac{1}{n}\sum_{t=1}^{n} RSSI^t, \quad \sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}\left(RSSI^t - \mu\right)^2}.$$

We choose the high probability zone as $P > 0.6$ [12], which is an experiment value. Through the above analysis, the select interval of RSSI is $[0.26\sigma + \mu,\ 3.09\sigma + \mu]$. Fig. 5 gives the PDF and probability distribution of one set of RSSIs' sampling value.

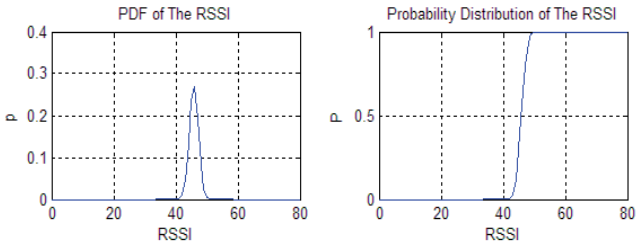

Fig. 5. RSSIs' probability distribution function (PDF).

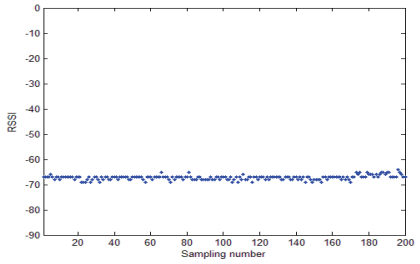

Fig. 6. RSSI Sampled values (2 meter).

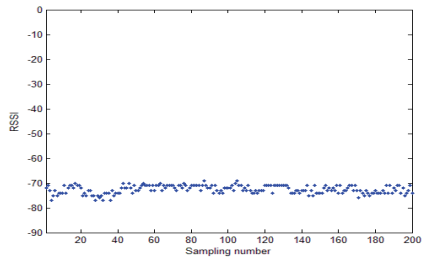

Fig. 7. RSSI Sampled values (3 meter).

Fig. 6 and Fig. 7 denote two sets of RSSI values for which the distances from tags to readers are 2 meters and 3 meters, respectively. The number n of RSSI values in each set is 200.

Above two kinds of circumstances each have two hundred set of values, these values are unfiltered data. Our filtering process is to eliminate those values which do not meet the interval $[0.26\sigma + \mu,\ 3.09\sigma + \mu]$. This interval is to use unfiltered data calculated out. After filtering, the rest of data is filtered data.

We calculate respectively the arithmetic mean, its estimated distance and estimation error of the two sets of values. Tab. 2 shows the result of the calculation. Tab. 2 shows the comparison of the estimated distance in two cases: filtered data and unfiltered data. From Tab. 2, we can see that the reliability of RSSI has improved significantly after Gaussian filtering.

Actual distance (meter)	Mean of RSSI (dBm)	Estimated distance (meter)	Estimation error
Unfiltered data			
2.0	-67.0073	1.744379	12.78%
3.0	-72.4887	2.656766	11.44%
Filtered data			
2.0	-68.2239	1.915108	4.24%
3.0	-73.7295	2.922223	2.59%

Tab. 2. Comparisons of two types of data.

2.3 Probability Model of Path Loss

We choose log-distance path loss model [13] as ideal free space propagation path loss formula:

$$PL(d) = PL(d_0) + 10n\log(d/d_0). \tag{2}$$

In (2), $PL(d)$ is propagation path loss, d is the actual distance from tags to readers, d_0 is reference distance and usually the value of d_0 is 1, n is the path loss exponent of the environment.

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. Hence equation (2) can be modified as:

$$PL(d) = PL(d_0) + 10n\log(d/d_0) + N(0,\sigma^2) \tag{3}$$

where $N(0, \sigma)$ denotes a Gaussian distribution with mean of 0 and variance of σ. Assume that P_r is the received power, P_t is the constant transmit power, we have

$$P_r = P_t - PL(d). \tag{4}$$

From the above, $d_0 = 1$, the expression $10n\log(d/d_0)$ in (3) changes to $10n\log(d)$, so (3) changes to (5)

$$P_r = P_t - PL(d_0) - 10n\log(d) - N(0,\sigma^2). \tag{5}$$

As we know, P_t - $PL(d_0)$ is a constant, so we can set it as a constant of α. $PL(d_0) = 10n \log(d_0)$, so $10n \log(d_0)$ is a part of constant α.

Then, equation (5) can be changed as:

$$P_r = \alpha - 10n \log d - N(0, \sigma^2) \quad (6)$$

where P_r is RSSI, it could be regarded as a Gaussian distribution with the mean of (α - 10$n$log$d$) and variance of σ^2. The PDF of RSSI is given by:

$$p(P_r = RSSI) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(RSSI - \alpha + 10n \log d)^2}{2\sigma^2}\right). \quad (7)$$

Through (6), we have

$$d_{estimate} = 10^{\frac{\alpha - RSSI}{10n}}. \quad (8)$$

Then equation (7) becomes equation (9), $d_{est} = d_{estimate}$

$$p(d_{est}) = \frac{10n}{\ln 10 d_{estimate} \sigma\sqrt{2\pi}} \exp\left[\frac{-\left(10n \log \frac{d_{actual}}{d_{estimate}}\right)^2}{2\sigma^2}\right]. \quad (9)$$

Fig. 8 depicts the PDF of the estimated distance. From Fig. 8 it can be concluded that:

1) If keeping variance σ^2 and actual distance d_{actual} unchanged, the curve becomes concentrated and steep with the path loss exponent n increasing.

2) If keeping variance σ^2 and path loss factor n unchanged, the dispersion and smoothness with the actual distance d_{actual} increase.

3) If keeping actual distance d_{actual} and path loss factor n unchanged, the curve shows that dispersion and smoothness with the variance σ^2 is increasing.

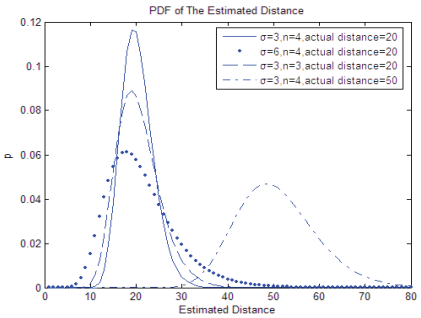

Fig. 8. PDF of the estimated distance.

2.4 Wheel Graph Model

In this section, we mainly discuss the positioning environment and the choice of location area. The actual positioning environment is too complex to use just three readers for large-scale indoor localization. So we divide a large area into many small square areas [14], as shown in Fig. 9. Each small area is covered by three readers and each reader's location is known. Tags will periodically broadcast the information to all readers around them, readers will collect information (all tags ID and RSSI) from tags and upload data (all readers ID, tags ID and RSSI) to the server. Then, the server processes the uploaded data. The data include three types of number, if the data can't be effectively paired, it is very easy to create chaos in the post-processing stage, i.e., we can't distinguish which RSSI is uploaded by which reader. We propose a new method, called wheel graph, to solve the problem of the positioning region selection.

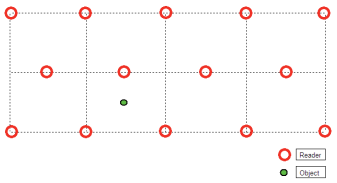

Fig. 9. Deployment of readers and tag.

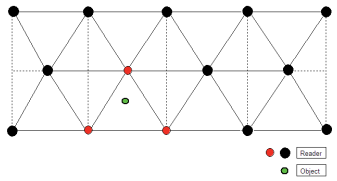

Fig. 10. Wheel network of environment.

From the previous section, we know that every reader will transmit RSSI information to the server. The server can estimate the distance from tags to readers using (8). $n1$ is the number of readers, the above situation can be described by the following formula, this formula can be expressed as:

$$\begin{bmatrix} Reader\_1 & Reader\_2 & Reader\_3 & Reader\_4 & \cdots & Reader\_n1 \\ d_1 & d_2 & d_3 & d_4 & & d_{n1} \end{bmatrix}.$$

After acquiring the distance information, we can establish the wheel graph model in our environment. Fig. 10 shows the wheel network of Fig. 9. As shown in Fig. 10, we connect all readers together. Each black or red dot represents a reader node.

Then, range $Dc = \{d_1\ d_2\ d_3\ d_4 \ldots d_5\}$ from small value to large one. And choose the smallest four values from it, and meet the condition of d_1 , d_2 , d_3 , $d_4 \in Dc$, given as

$$\begin{bmatrix} Reader\_1w & Reader\_2w & Reader\_3w & Reader\_4w \\ d_{1w} & d_{2w} & d_{3w} & d_{4w} \end{bmatrix},$$

and $d_{1w} \le d_{2w} \le d_{3w} \le d_{4w}$. The following are the absolute coordinates of the four readers,

$$Reader\_1w = (x_{1w}, y_{1w}),\ Reader\_2w = (x_{2w}, y_{2w}),$$
$$Reader\_3w = (x_{3w}, y_{3w}),\ Reader\_4w = (x_{4w}, y_{4w}).$$

We can utilize the two shortest distances, d_{1w} and d_{2w} , to determine two closest readers near the object. Then we regard the two readers as two vertices of the triangle. Next, we use *Reader*_3*w* and *Reader*_4*w* respectively as the third vertex to form a triangle, and judged by using the following rules (Low of Cosines).

$$\cos\theta_1 = \frac{(x_{2w}-x_{1w})(x_{3w}-x_{1w})+(y_{2w}-y_{1w})(y_{3w}-y_{1w})}{\sqrt{(x_{1w}-x_{3w})^2+(y_{1w}-y_{3w})^2}\sqrt{(x_{1w}-x_{2w})^2+(y_{1w}-y_{2w})^2}}$$

$$\cos\theta_2 = \frac{(x_{1w}-x_{2w})(x_{4w}-x_{2w})+(y_{1w}-y_{2w})(y_{4w}-y_{2w})}{\sqrt{(x_{2w}-x_{4w})^2+(y_{2w}-y_{4w})^2}\sqrt{(x_{1w}-x_{2w})^2+(y_{1w}-y_{2w})^2}}$$

When above formulas simultaneously satisfy the following conditions: $\cos\theta_1 > 0$ & $\cos\theta_2 > 0$, we take the point as the third vertex. Using wheel graph method, we can determine the small area covered by three readers. And our positioning region is a square area containing the triangle, while the horizontal side of the triangle is the square area's one side. Fig. 11 displays the wheel graph of the whole process. Fig. 12 shows the small square positioning area.

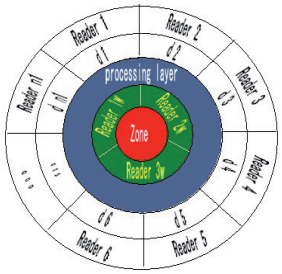

Fig. 11. Wheel graph of the whole process.

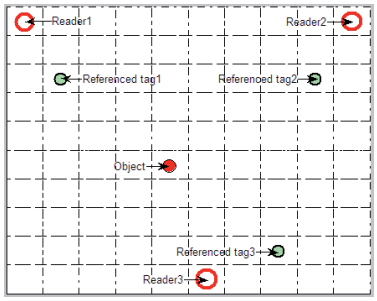

Fig. 12. Positioning area.

2.5 Bayesian Localization Scheme

2.5.1 Calibration of the Propagation Parameters

In the previous section, we have determined the positioning area. In order to calculate the probability in (7) or (9), and the estimated distance in (8), we need to know the propagation parameters α, n and σ. In our Bayesian localization algorithm, the on site reference tags and object tags are used to calibrate the propagation parameters during the process of localization. Every reader has a reference tag deployed in the environment within its range. The reference tags are static with known locations. After detection, the readers report the values of RSSIs of tags to the server; the server will process the received data and calibrate the propagation parameters. In the following part, we will explain this process in detail.

The actual situation is shown in Fig. 12. There are three readers in the specific area, reader_1, reader_2 and reader_3. There are four tags in the space, three reference tags and one object tag. Suppose that during a period of time, time slot 1 to time slot w, received signal strength recorded as $\{RSSI^t_j(i)\}$, j = 1,2,3, i = 1,2,3,4, t = 1,2,3…w. Here, the $\{RSSI^t_j(i)\}$ denotes reader_j report the RSSIs coming from tag_i during time slot 1 to time slot w. We can calculate the average value of RSSIs, $\overline{RSSI_j(i)}$ denoted as:

$$\overline{RSSI_j(i)} = \alpha_j - 10n_j \log(dist(i,j)). \tag{10}$$

From (6) and (10), we can get the following formula.

$$RSSI_j(i) = \overline{P_r(dist(i,j))}. \tag{11}$$

First, we calibrate the path loss exponent n_j. We can derive equation (12), (13) from (10) and (11). $dist(i,j)$ means the distance from tag_i to reader_j. We have:

$$\overline{RSSI_j(j)} = a_j - 10n_j log(dist(j,j)), \tag{12}$$

$$\overline{RSSI_j(i)} = a_j - 10n_j log(dist(i,j)). \tag{13}$$

Then the difference of two equations, (12) and (13) is

$$\overline{RSSI_j(i)} - \overline{RSSI_j(j)} = 10n_j \log\frac{dist(i,j)}{dist(j,j)}. \tag{14}$$

Then we can get the path loss exponent $n_j(i)$, which denotes one calibration from reader_j to tag_i, and $i \neq j$. In this section, j = 1,2,3, and i = 1,2,3.

$$n_j(\mathrm{i}) = \frac{\overline{RSSI_j(\mathrm{i})} - \overline{RSSI_j(\mathrm{j})}}{10\log\dfrac{dist(i,j)}{dist(j,j)}}. \tag{15}$$

For reader_j, the path loss factor is calculated twice, we use its average value in our system.

$$n_j = \sum_{i=1,2,3}\frac{n_j(i)}{2}. \tag{16}$$

n_j denotes reader_j's path loss factor to the environment. Because there are three readers in our space, the path loss factor should be calculated three times. Secondly, we calibrate the α_j, from (10), we have

$$\alpha_j(\mathrm{i}) = \overline{RSSI_j(i)} + 10n_j log(dist(i,j)), \tag{17}$$

$$\alpha_j = \sum_{i=1,2,3}\frac{a_j(i)}{2}, \quad (i \neq j). \tag{18}$$

Finally, we calibrate the value of σ, as in formula (19). N_r and N_t are the numbers of readers and tags. In our preset situation, N_r = 3, N_t = 3, and $std(RSSI_j(i))$ denotes the standard deviation for vector $[RSSI^1_j(i), RSSI^2_j(i),\ldots, RSSI^w_j(i)]$.

$$\sigma = \sqrt{\frac{1}{N_r N_t}\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}[std(\mathrm{RSSI}_j(i))]}. \tag{19}$$

2.5.2 Probability Localization Algorithm

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. According to the Bayesian inference [15], we can obtain the following formula.

$$p\left(x|RSSI^t\right)=\frac{p\left(RSSI^t|x\right)p\left(x\right)}{\int p\left(RSSI^t|x\right)p\left(x\right)dx}=\frac{p(RSSI^t\,|\,x)p(x)}{p(RSSI^t)} \quad (20)$$

where $p(x)$ is the priori probability of the object's location. $p(RSSI^t|x)$ can be derived from (7), which means the PDF of RSSI under the situation of location information 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 equation (20), we have following formula [16]

$$p(x\,|\,RSSI^t)\propto p(RSSI^t\,|\,x)p(x)\,.$$

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

$$Bel(x)=\beta\,p\left(RSSI^t|x_k\right)Bel^-(x). \quad (21)$$

$Bel^-(x)$ can be computed from equation (22).

$$Bel^-(x)=\int p\left(x_k|x_{k-1}\right)p\left(x_{k-1}\middle|RSSI^t\right)dx_{k-1} \quad (22)$$

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.

$$p(x_k\,|\,x_{k-1})=\begin{cases}1 & x_k\neq x_{k-1}\\ 0 & x_k=x_{k-1}\end{cases} \quad (23)$$

For a static system, equation (22) can be converted into equation (24).

$$Bel^-(x_k)=Bel(x_{k-1}). \quad (24)$$

Then we can get our Bayesian estimation model.

$$Bel(x_k)=\beta\,p\left(RSSI^t|x_k\right)Bel(x_{k-1}). \quad (25)$$

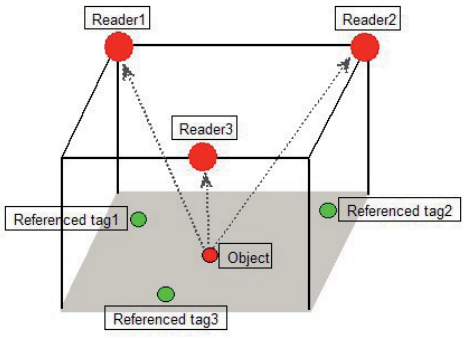

Fig. 13. Deployment environment.

As shown in Fig. 12, we divide the location area into N grids. In our simulation, the grid size is 40 cm × 40 cm, 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 $C_{(gxi,gyi)}$, $i\in(1,2,3\ldots N)$. In this area, small dots represent tags, green ones are inference tags and red one is object tag. Three big red circles are readers. A 2-dimensional or 3-dimensional map can be shown in Fig. 12 and Fig. 13 respectively.

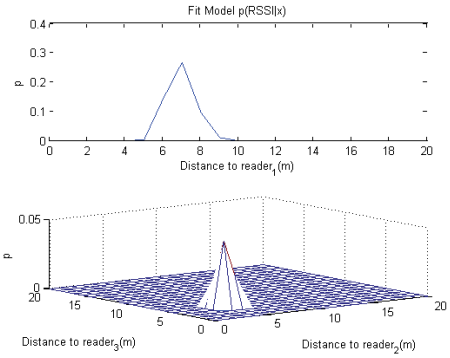

Fig. 14. Fit Model of one tracked tag.

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. 14 shows the Fit Model of a tracked point.

$$p(RSSI^t\,|\,x_k)=\prod_{j=1}^{3}p(RSSI_j^t\,|\,x_k)\,. \quad (26)$$

There are two pictures in Fig. 14, the precondition of this model is that RSSIs 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. 14, we can know that with the Fit Model relatively large values are achieved, when the estimated locations are close to the actual location. The maximum value is obtained when the estimated location is actual location.

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

$$Bel'(x_k)=\frac{Bel(x_k)}{\max\left(Bel\left(x_k=(x_k,y_k)\right)\right)}. \quad (27)$$

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. We can achieve one more precise estimated area, if increasing the recursion times R.

3. System Performance Analysis

This section is divided into three parts. First we discuss the average positioning distance error. Then we will discuss the performance of the system with changing parameters. At last, we evaluate the system performance. As we know, the performance will be affected by several undetermined parameters, for example, recursion time R, window size w and threshold T.

3.1 Average Localization Distance Error

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

$$ALE = \frac{\sum D\left(C_{(gxk,gyk)}\right) Bel'\left(x_k = C_{(gxk,gyk)}\right)}{\sum Bel'\left(x_k = C_{(gxk,gyk)}\right)}. \quad (28)$$

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

$$D\left(C_{(gxk,gyk)}\right) = \sqrt{\left(gx_k - x_0\right)^2 + \left(gy_k - y_0\right)^2}. \quad (29)$$

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

3.2 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 localization distance error (*ALE*): filtered data and unfiltered data, recursion time R, threshold T, path loss exponent n, standard deviation σ, and window size w, etc.

3.2.1 Impact of Filtered Data, Unfiltered Data and Recursion Time

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 = 40 cm, threshold $T = 0.5$, path loss exponent $n = 3$, standard deviation $\sigma = 1.45$, $\alpha = -59.7581$, window size $w = 200$.

Fig. 15 depicts *ALE* as a function of the recursion time for the two scenarios. In the first scenario, we observe that the decrease of *ALE*'s rate is very slow as we increase the recursion time. After several rounds of recursion, *ALE* is still large. But in the second scenario, we observe that *ALE* decreases after increasing the recursion time. This is due to the use of the Gaussian Filter which significantly improves the reliability of RSSI values.

Fig. 16 illustrates *ALE* as a function of the recursion time for different objects in the above second scenarios. From Fig. 16, we observe that *ALE* 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.

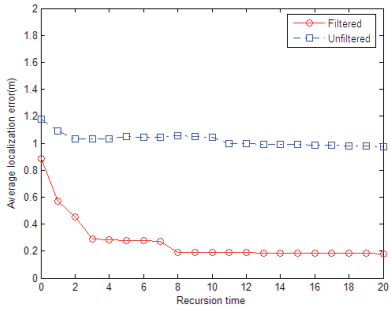

Fig. 15. Impact of Gaussian filter.

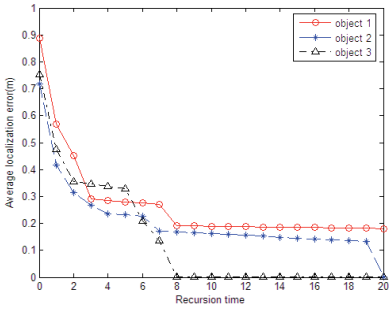

Fig. 16. Tracked three objects.

3.2.2 Impact of Window Size

Experiment 2: The following system is configured as: grid size = 40 cm, threshold $T = 0.5$, path loss exponent $n = 3$, standard deviation $\sigma = 1.45$, $\alpha = -59.7581$. Fig. 17 plots *ALE* 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. 17.

From Fig. 17, we observe that *ALE* decreases as the window size w increases. In other words, the greater the window size w, the higher the localization accuracy. We also find that the *ALE* decrease rate depends on the window size. The *ALE* decrease rate will be very high when the window size is relatively large.

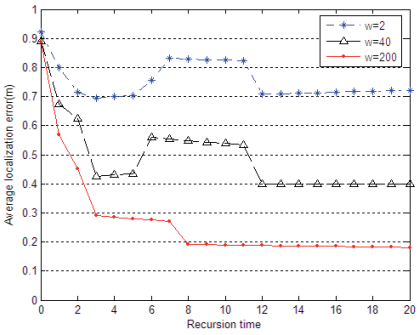

Fig. 17. Impact of window size.

3.2.3 Impact of Threshold, Path Loss Exponent and Standard Deviation

Experiment 3: The system is configured as: grid size = 40 cm, path loss exponent $n = 3$, standard deviation $\sigma = 1.45$, $\alpha = -59.7581$. 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. 18 plots *ALE* as a function of the threshold T.

From Fig. 18, it implies zero error for $T = 1$, this is true only if no tags are found, i.e., with high threshold, the error is low but almost no tags are identified. So, we must have a principle, tags should be identified at the current threshold, if no tags are found at the current threshold, we discard the value.

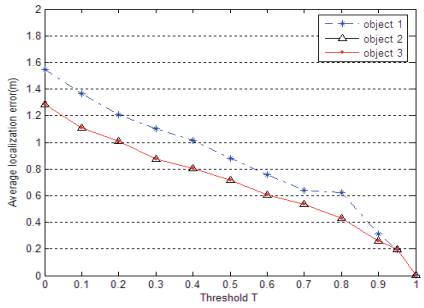

Fig. 18. Impact of threshold.

Experiment 4: The system is configured as: grid size = 40 cm, threshold $T = 0.5$, standard deviation $\sigma = 1.45$, $\alpha = -59.7581$. 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. 19 plots *ALE* as a function of the path loss exponent n.

From (15) and (16), we can know that the path loss exponent is not constant for the environment but a function of both reader and tag position.

From Fig. 19, we can see that localization accuracy increases with the increase of path loss exponent n. In practical applications, the path loss exponent shouldn't be set manually but be set according to the environment. When the environment path loss exponent is bigger, the system positioning accuracy is higher.

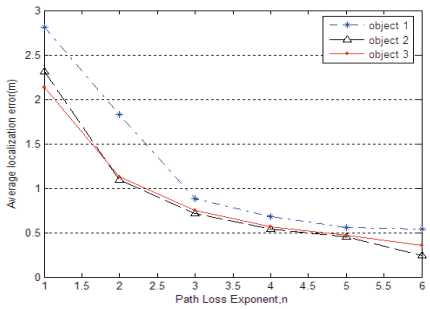

Fig. 19. Impact of path loss exponent.

Experiment 5: The system configuration: grid size = 40 cm, path loss exponent $n = 3$, threshold $T = 0.5$, $\alpha = -59.7581$. In this part, we evaluate the impact of standard deviation on the localization accuracy. Fig. 20 plots *ALE* as a function of standard deviation σ.

From Fig. 20, we can see that localization accuracy decreases with standard deviation σ increasing. The localization accuracy with small standard deviation is better.

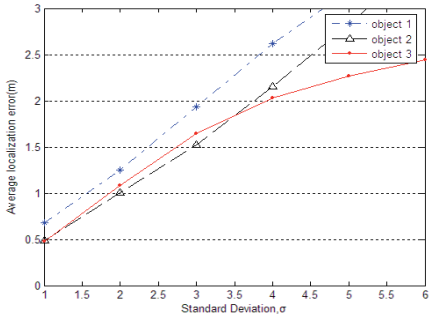

Fig. 20. Impact of standard deviation.

3.3 System Performance

In this section, we test the system performance using the following experiment.

Experiment 6: This experiment is designed to probabilistically evaluate the system localization performance under the following settings: Using the same reader network topology as shown in Fig. 12, we track 169 objects which are uniformly distributed within the basic detection area. The other simulation parameters are: grid size = 40 cm, path loss exponent $n = 3$, standard deviation $\sigma = 1.45$, $\alpha = -59.7581$, 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. 21 shows the cumulative Average Error Distance distribution obtained from scenarios 1 and 2. We notice that in over 90% percent of the objects the localization *ALE* falls from within 50 cm. As expected, the increasing recursion time R will lead to higher localization accuracy. Furthermore, from Fig. 18, we also can get the conclusion that in most of the situation, localization accuracy with large T is better.

4. Conclusion

In this paper, we presented a new method, an easy-setup and cost-effective indoor localization method based on off-the-shelf active RFID technology combined with GF, WGM and PLA. Our system relies on a hierarchical architecture to cover an indoor environment. The proposed approach can calibrate the propagation parameters according to the environment, reduce the uncertainty of localization, and at the same time obtain high positioning accuracy under conditions of large-scale indoor positioning. Using MATLAB simulations we have evaluated the performance of our proposal. In the destined system settings, the simulation results show that in 90% percent of the loca-

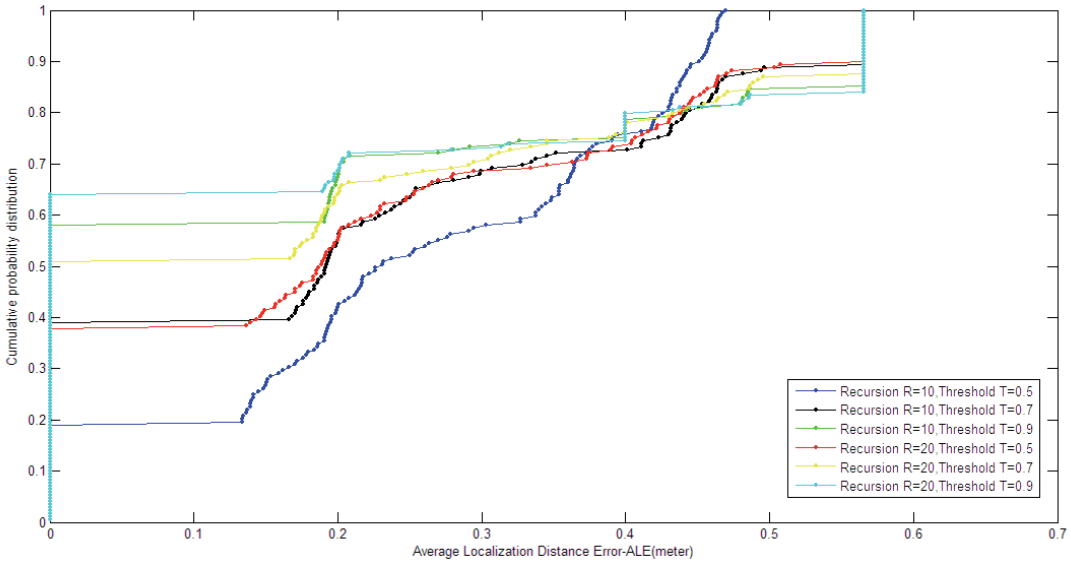

Fig. 21. 169 objects' cumulative probability distribution.

lization estimation, the system provides objects location with the *ALE* less than 50 cm, and in 70% percent of the localization estimation, the system provides objects location with the *ALE* less than 20 cm. The simulation also shows that the system performance improves with the higher values of recursion time, window size, path loss exponent and with the lower propagation standard deviation. The simulation results can prove that the proposed system is an accuracy and cost-effective candidate for indoor localization.

Acknowledgements

Authors are grateful to the anonymous reviewers for their helpful comments. They also wish to note that the research work is supported by the National Natural Science Foundation of China (NO.61101209) and Shanghai Normal University key project (A7001-12-002006).

References

[1] *Garmin Corporation*, *About GPS*. [Online]. Available at: http://www.garmin.com/aboutGPS/.

[2] HUI, L., HOUSHANG, D., PAT, B., JING, L. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics- Part C*, 2007, vol. 37, no. 6, p. 1067-1080.

[3] HIGHTOWER, J., BORRIELLO, G. A survey and taxonomy of location sensing systems for ubiquitous computing. *CSE 01-08-03*, University of Washington, Department of Computer Science and Engineering, Seattle (USA). Technical Report, 2001.

[4] GUOQIANG, M., BARIS, F., BRIAN, D. O. A. Wireless sensor network localization techniques. *Computer Networks*, 2007, vol. 51, no. 10, p. 2529–2553.

[5] *Radio Frequency Identification (RFID) home page*. [Online]. Available at: http://www.aimglobal.org/technologies/rfid/.

[6] YE, Y., JUN, Z., JIAN, Y. Design of World Expo tour sites guide system based on RFID technology. In *Proceedings of 2010 International Conference on Audio Language and Image Processing (ICALIP 2010)*. Shanghai (China), 2010, p. 1026-1030.

[7] HIGHTOWER, J., WANT, R., BORRIELLO, G. SpotON: An indoor 3d location sensing technology based on RF signal strength. *UW-CSE 00-02-02*, University of Washington, Department of Computer Science and Engineering, Seattle (USA). Thesis, 2000.

[8] NI, L., LIU, Y., LAU, Y. C., PATIL, A. LANDMARC: Indoor location sensing using active RFID. In *Proceedings of the 1st IEEE International Conference on Pervasive Computing and Communications (PERCOM'03)*. Dallas (USA), 2003, p. 407-415.

[9] MAJA, S., MLADEN, R., DINKO, B. RF Localization in indoor environment. *Radioengineering*, 2012, vol. 21, no. 2, p. 557–567.

[10] ZEPERNICK, H. J., WYSOCKI, T. A. Multipath channel parameters for the indoor radio at 2.4 GHz ISMband. In *Proceedings of Vehicular Technology Conference*, *1999 IEEE 49th*. Houston (USA), 1999, vol. 1, p. 190-193.

[11] ITO, K. Gaussian filter for nonlinear filtering problems. In *Proceedings of the 39th IEEE Conference on Decision and Control*. 2000, vol. 2, p. 1218-1223.

[12] MINGHUI, Z., HUIQING, Z. Research on model of indoor distance measurement based on receiving signal strength. In *Proc. of Internat. Conf. on Computer Design and Applications (ICCDA 2010)*. Qinhuangdao (China), 2010, vol. 5, p. 54-58.

[13] RAPPAPORT, T. S. *Wireless Communications Principles and Practices*. Prentice-Hall Inc, 2002.

[14] WOOYONG, L., KYEONG, H., TAEYOUNG, K., DOODEOP, E., JONGOK, K. Large scale indoor localization system based on wireless sensor networks for ubiquitous computing. *Wireless Personal Communications*, 2012, vol. 63, no. 1, p. 241-260.

[15] MADIGAN, D., ELNAHRAWY, E., MARTIN, R. Bayesian indoor positioning systems. In *Proceedings of 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005)*. Miami (USA), 2005, vol. 2, p. 1217-1227.

[16] CHRISTOPHER, M. B. *Pattern Recognition and Machine Learning*. New York: Springer-Verlag Inc, 2006.

About Authors ...

Long XIAO was born in 1988 in Wuhan, China. He received his B.S. degree in Electronic and Information Engineering from the Huanggang Normal University in 2011. He is now a graduate at Shanghai Normal University, China. His research interests include wireless positioning, wireless sensor network and signal processing. He has over four pending patents.

Ye YIN was born in 1961 in Jiangsu province, China. He received his M.S. and Ph.D degrees from the Tongji University. He is an associate professor of the College of Information, Mechanical and Electrical Engineering. His research interests include relativity theory, information theory, gravitational field communication, voice communication code.

Xiangnong WU received her B.S. and M.S. degrees from the Huazhong University of Science and Technology, China. She received Ph.D degree from the Nanyang Technological University, Singapore. She is an associate professor of the College of Information, Mechanical and Electrical Engineering. Her research interests include optical fiber communication.

Jianwei WANG received his M.S. degrees from the Tongji University. His research interests include RFID and signal processing. He is the founder of Shanghai Zhen Zhuo Electrical Technology Ltd. Co.