An Improved Indoor Geolocation Algorithm based on Log-distance Path Loss Model

Pei Liang[1,2 a], Liu Shixuan,[1,2,b], Wang Bo[1,2,c], Li Wenqing[1,2,d]

[1]Shandong Provincial Key Lab of Ocean Environment Monitoring Technology, Qingdao, China

[2]Institute of Oceanographic Instrumentation, Shandong Academy of Science, Qingdao, China

[a]peiliang2002@163.com, [b] 321_go@sohu.com, [c] bob80.wang@hotmail.com, [d] livenson@163.com

Keywords: Indoor geolocation, RSSI, fingerprint, log-distance path loss model.

Abstract. This paper proposes an indoor localization algorithm based on Log-distance path loss model. The two parameters of the Log-distance path loss model are estimated by the measured RSSI data. Then the two step filtering was carried out to keep the RSSI signals with the higher probabilities. Finally the Taylor's series iteration method is used to compute the position coordinate of the moving objects. The experiment results show that the location error is no more than 3m.

Introduction

In recent years, the application of the location based services has been related to almost every aspect in our daily life. GPS and Assisted GPS can achieve satisfactory location accuracy in outdoor applications. However, GPS and A-GPS can't be applied to the indoor geolocation because GPS signals indoor are too weak to be detected. To the indoor geolocation, the wireless local area network (WLAN) based systems maybe the best solution [1,2]. It's convenient to use the existed WLAN network.

In the indoor geolocation, we can use three methods, that is, TOA, AOA, and RSSI [3-7]. The TOA and AOA methods need extra hardware to realize the location purpose, that is, the TOA method needs the synchronization precision of ns, and the AOA method need antennas. Moreover, the problem of multi-path effect is normal in indoor environments, sometimes the direct path is obstructed completely, which is the NLOS problem and will make the precise location impossible for TOA and AOA. And the coverage of AP is about 100m, so the propagation time of the wireless signals of TOA is difficult to be measured. So the RSSI method is adopted in this paper.

The RSSI log-distance path loss model is developed based on sufficient actual RSSI measurements in indoor environments. Then the improved log model is used to filter the RSSI signals. And the Taylor's series iteration method is adopted to realize location estimation of the testing points. Testing results show that the algorithm developed in this paper can achieve location error of 3m.

RSSI Log-distance path loss model

In WLAN based indoor geolocation technology, there are two location method can be adopted, that is, fingerprinting method and radio propagation modeling method. And the fingerprinting method is mostly used, but the fingerprinting data is closely related to the indoor environment, when the environment changes such as a small difference in the layout of the furniture will result to great difference in the fingerprinting data. The radio propagation modeling method also is sensitive to the environment. In order to improve the location precision, an improved radio propagation modeling method is developed in this paper.

The radio propagation modeling method uses path loss model to estimate the distance. And the path loss model often used include Free Space Propagation Model, Log-Distance Distribution Model, and Log-Distance Path Loss Model [8-10]。

To the indoor geolocation, the environment is complex, and the factors such as the building structure and the layout of the objects indoor will increase the path loss. Therefore, the Log-Distance Path Loss Model is adopted in this paper.

Log-Distance Path Loss Model. The Log-Distance Distribution mode is described by formula(1)：

$$PL(d) = PL(d_0) + 10n\lg(\frac{d}{d_0}) + X_s \tag{1}$$

In formula (1), d is the distance between the transmitter and the receiver, $PL(d)$ is the path loss in dB, d_0 is reference distance, X_s is a zero mean variable with the variance s^2, and the RSSI measured comply with formula(2)：

$$RSSI = P_{sent} - PL(d) \tag{2}$$

In formula(2)，RSSI is the signal strength received at the testing points in dB, P_{sent} is the transmit power in dB. If the RSSI is A when $d = d_0$, then we have $A = P_{sent} - PL(d_0)$, then we have:

$$PL(d_0) = P_{sent} - A \tag{3}$$

Then formula (1) can be described by formula (4):

$$PL(d) = P_{sent} - A + 10n\lg(\frac{d}{d_0}) + X_s \tag{4}$$

And the reference distance $d_0 = 1\text{m}$，then using formula (4) to represent $PL(d)$ in formula (2), we have:

$$RSSI = A - 10n\lg(\frac{d}{d_0}) - X_s \tag{5}$$

And here X_s is a zero mean variable, so we have:

$$\overline{RSSI} = A - 10n\lg(d) \tag{6}$$

And $\overline{RSSI}$ is the average of RSSI based on many testing results. From formula (6),we have:

$$d = 10^{\frac{A - \overline{RSSI}}{10n}} \tag{7}$$

So formula (7) is adopted to estimate the distance between the transmitter and the moving objects. In formula (7), two parameters A and n should be decided, and these two parameters will be estimated in the following section.

Parameter estimation. In the indoor geolocation, we need three given nodes to decide the coordinate of the object. In fig.1 node 5 is the testing node needed to be located, and node 1, 2, 3 are the three given nodes with the precise coordinates, and these three is near to node 5 with the minimum distance. The three given nodes can communicate to each other. And d_{12}，d_{13}，d_{23} are the distance between the three given nodes. Here, the unknown node 5 and the three given nodes are very near to each other, so the wireless channel can be taken as the same. So we can use the three given nodes to estimate the parameters in formula (7).

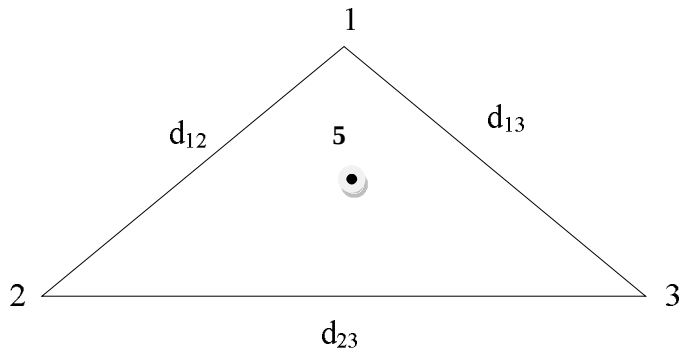

Fig.1 basic idea of parameter estimation

Supposing node 1 is the transmitter, the RSSI received at node2 and node3 is R_2 and R_3, and according to $RSSI = A - 10n\lg(d)$, we can have:

$$\begin{cases} R_2 = A_1 - 10n_1 \lg d_{12} \\ R_3 = A_1 - 10n_1 \lg d_{13} \end{cases} \tag{8}$$

Then the two parameters A_1 and n_1 of node1 can be solved from equation (8):

$$\begin{bmatrix} A_1 \\ n_1 \end{bmatrix} = \begin{bmatrix} 1 & -10\lg d_{12} \\ 1 & -10\lg d_{13} \end{bmatrix}^{-1} \begin{bmatrix} R_2 \\ R_3 \end{bmatrix} \tag{9}$$

Similarly, the parameters of node2 and node3 can be achieved:

$$\begin{bmatrix} A_2 \\ n_2 \end{bmatrix} = \begin{bmatrix} 1 & -10\lg d_{21} \\ 1 & -10\lg d_{23} \end{bmatrix}^{-1} \begin{bmatrix} R_1 \\ R_3 \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} A_3 \\ n_3 \end{bmatrix} = \begin{bmatrix} 1 & -10\lg d_{31} \\ 1 & -10\lg d_{32} \end{bmatrix}^{-1} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \tag{11}$$

Then we use the average of node1, node2 and node3 to represent the parameters of the triangle area described by the three nodes:

$$\begin{cases} A = \dfrac{A_1 + A_2 + A_3}{3} \\ n = \dfrac{n_1 + n_2 + n_3}{3} \end{cases} \tag{12}$$

Based on above, we can use formula (7) to estimate the distance between the given AP and the moving object.

RSSI filtering

Distribution of RSSI. The indoor RSSI is sensitive to the surrounding environments because of multipath effect, such as the layout of the furniture and the people moving around, Fig. 2 is the RSSI measured indoor. We can see in Fig.2 that the RSSI doesn't comply with the Log-Distance Path Loss Model well with fast fading, and the RSSI doesn't comply with distance, and our testing show that RSSI in a fixed indoor position comply with the normal distribution as shown in Fig.2.

So before we use formula (7), the RSSI signals must be processed using some statistical methods. And we will use two steps of filtering to obtain the RSSI signals with the maximum probability, so high precision of positioning can be achieved.

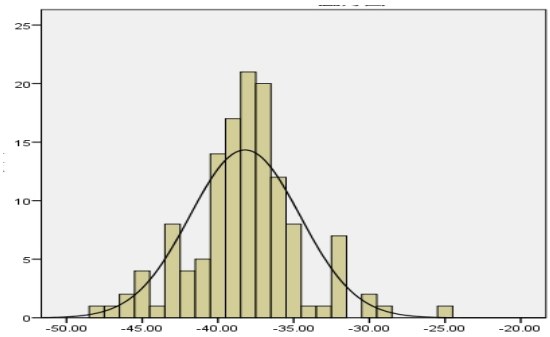

Fig.2 Distribution of RSSI

RSSI filter. Average model and Gaussian model are often used to process the RSSI signals [14]. The average model simply uses the average of the RSSIs by the multiple measurements, and this will lead to small probability error, so we can't achieve precise position estimation. To the Gaussian model, the average of signals with higher probabilities filtered out by the Gaussian model will be adopted to estimate the object position. But in paper [17] it is concluded that the indoor RSSI signals have positive bias because the NLOS problem is normal in indoor environments, so the Gaussian model is also not accurate for indoor geolocation.

An improved Log-Distance Path Loss Model is developed in this paper, in which two steps of filtering is adopted. The first filtering use the standard deviation σ to filter the RSSIs with big errors, the second filtering is based on the Log-Distance model, so we can keep the RSSIs with higher probabilities.

In the first step, to the Gaussian RSSI signals, the probability of the RSSIs in the range (m - s , $m + s$)is 0.6526, so we can use s to keep the RSSIs with higher probabilities.

Gaussian distribution function is formula (13):

$$F(x) = \frac{1}{s\sqrt{2p}} e^{\frac{(m-x)^2}{2s^2}} \tag{13}$$

The average is:

$$m = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{14}$$

And variance:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - m)^2 \tag{15}$$

So we can use the RSSIs to estimate m and s , then keep the RSSIs in the range of $(m-s, m+s)$, in which the RSSIs with big errors are mostly removed. And this is the first filtering, in which every RSSI measurement is independent, so the filtered RSSIs still comply with Gaussian distribution.

In the first filtering, the RSSIs with big errors are removed, so in the second filtering, we will use the Log-distance model to further process the RSSIs.

The distribution of Logarithmic normal distribution is:

$$f(x) = \frac{1}{-xs_1\sqrt{2p}} e^{-\frac{(\ln(-x)-m_1)^2}{2s_1^2}} \tag{16}$$

The average is:

$$m_1 = \frac{\sum_{k=1}^{n}\ln(-x_k)}{n} \tag{17}$$

And variance:

$$s_1^{\ 2} = \frac{\sum_{k=1}^{n}(\ln(-x_k) - m_1)^2}{n} \tag{18}$$

Then we will use the range decided by formula (19) to filter the RSSIs with high probability, and here the parameter 0.6 is decided by large scale experiments.

$$\frac{0.6}{m_1 s_1 \sqrt{2p}} \le f(x) \le \frac{1}{m_1 s_1 \sqrt{2p}} \tag{19}$$

Then we can use the RSSIs remained by the two step filtering to estimate the RSSI of the testing node:

$$m1 = \frac{\sum_{k=1}^{n} \ln(-x'_k)}{n} \tag{20}$$

$$\bar{x} = -e^{m1} \tag{21}$$

Here, x'_k is the the remaining RSSI signals, and n is the number of x'_k, $\bar{x}$ is the estimated RSSI of the testing point. So we can use formula (12) to estimate the parameters in formula (7), then use the $\bar{x}$ to estimate the distance between the transmitter or AP and the moving object, and three APs are needed to decide the accurate coordinate of the object.

Positioning algorithm

Trilateration and triangulation are often used to calculate the coordination of the object, but these two methods can't realize accurate coordinate estimation in indoor environments. We will use Taylor's series iteration method to estimate the coordinate.

The idea of Taylor's series iteration method is to expand the positioning function at the testing node, ignore the second item and items above second, then the bias will be used to upgrade the coordinate estimation, with the iteration step by step, finally the coordinate estimated will be convergent to the real position of the testing point.

Supposing the real coordinate of the object is (x, y), the coordinate of the given nodes are (x_i, y_i), N is the number of the given nodes, m_i is the parameter of the i_{th} given node, m_i is the real value of m_i, e_i is the measurement error obeys zero mean normal distribution with the covariance matrix R. Here we use f_i to represent the coordinate of the testing node:

$$f_i(x, y, x_i, y_i) = m_i = m_i - e_i, i = 1,2,...,N \tag{22}$$

The steps of Taylor's series iteration method are described by the following:

Step 1. Supposing the initial position in the iteration is (x_v, y_v), the estimation error is d_x, d_y, then we have:

$$x = x_v + d_x, y = y_v + d_y \tag{23}$$

Step 2. Expand f_i at (x_v, y_v) using Taylor series, and ignore the nonlinear items, we have:

$$f_{iv} + a_{i1}d_x + a_{i2}d_y \approx m_i - e_i, i = 1,2,...,N \tag{24}$$

here $f_{iv} = f_i(x_v, y_v, x_i, y_i), a_{i1} = \frac{\partial f_i}{\partial x}\Big|_{x_v, y_v}, a_{i2} = \frac{\partial f_i}{\partial y}\Big|_{x_v, y_v}$ $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2j} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} \end{pmatrix}$

And if the number of the unknown variables and the observation equations are the same, and matrix A is full rank, then the weighted least square estimation of d is:

$$d = (A^T R^{-1} A)^{-1} A^{-T} R^{-1} z \tag{25}$$

Step 3. If $e < e_0$, then exit, we can have the position estimation as: $\hat{x} = x_v + d_x$, $\hat{y} = y_v + d_y$ Else, go to step (4);

Step 4. Update x, y using $x = x_v + d_x, y = y_v + d_y$,then go to step(2).

Here we can use the following equations to decide the value of ε:

$$e = \sqrt{d_x^{\ 2}(n+1) + d_y^{\ 2}(n+1)} \tag{26}$$

$$e = \left| \sqrt{d_x^{\ 2}(n) + d_y^{\ 2}(n)} - \sqrt{d_x^{\ 2}(n+1) + d_y^{\ 2}(n+1)} \right| \tag{27}$$

$$e = \max(|d_x(n+1)|, |d_y(n+1)|) \tag{28}$$

$$e = \max(|d_x(n) - d_x(n+1)|, |d_y(n) - d_y(n+1)|) \tag{29}$$

Here n is the number of iterations, and we should decide the initial coordinate near the actual position by prior knowledge.

Experiments

Experiment 1. The first experiment is for the two step filtering. We get 100 RSSI samples when d=1m and 200 RSSI samples when d=5m, then we use the two step filtering method to filter the RSSI samples. In the first filtering, we estimate μ and σ firstly, then keep the RSSI samples in $(m - s, m + s)$ with the probability 0.6526. In the second filtering, we use the Log-distance model, that is, formula (19) to filter the RSSI samples.

Fig.3 is the original RSSI samples as *d*=1m, it's clear that the value of RSSI samples is not stable fluctuating in the range [-28dB,-44dB]. Fig.4 is the result of the first filtering, the fluctuating is limited to [-36dB,-40dB], so the samples with big errors are removed. And Fig.5 is the result of the second filtering, we can see that the amplitude of the RSSI signals is fluctuating in a small scale around the real value.

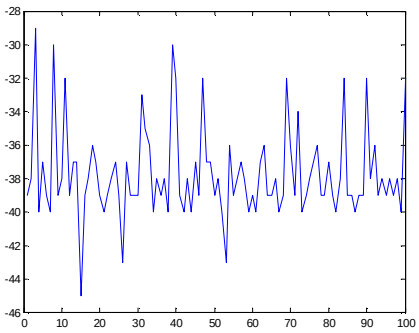

Fig.3 Unfiltered RSSI

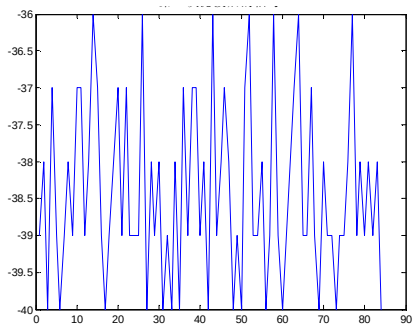

Fig.4 RSSI after the first filter

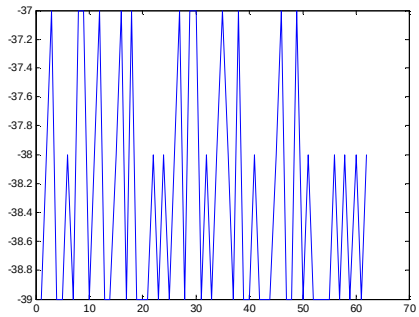

Fig.5 RSSI after the second filter

And Fig.6 is the original RSSI samples as d=5m, Fig.7 and Fig. 8 are results of the two step filtering, we keep the RSSI samples with higher probability. So we can use the RSSI samples remained to estimate the average.

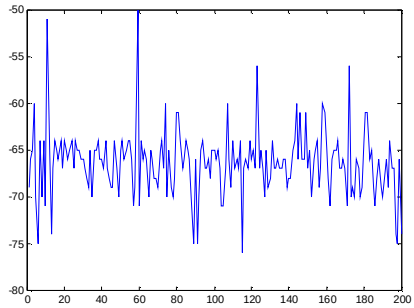

Fig.6 Unfiltered RSSI

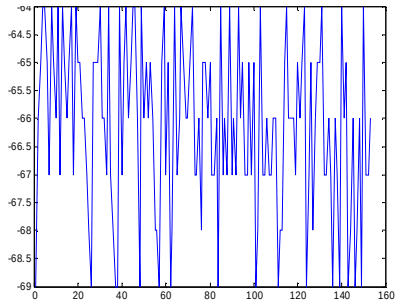

Fig.7 RSSI after the first filter

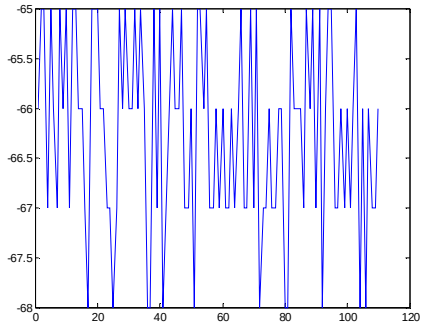

Fig.8 RSSI after the second filter

Experiment 2. In the second experiment, we use the Taylor series method to calculate the object position. The testing area is a lab with the size 9.5m×6m, and we choose 25 testing nodes as shown in Fig. 4.9, in which x axis is the length and y axis is the width of the lab. The coordinate of the three given Aps is AP1:(1.71，2.4),AP2:（7.79，2.4）,AP3:（2.09，4.8）.So we can use formula (9)~(12) to estimate the model parameter A and n, and we get A=48， n=2.9.

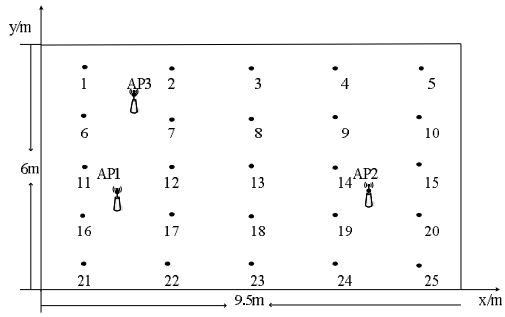

Fig.9 Layout of the indoor test points

To the every testing node, 100 samples are adopted to estimate the RSSI average using the two step filtering, then the RSSI average can be estimated. Finally the Taylor's series iteration method can be used to estimate the coordinate of the object. Fig. 10 is the position estimation error of the 25 testing nodes. We can see that the estimation error fluctuating around 2m with the maximum 3m. This is superior to the traditional trilateration and triangulation method, which is in the range [3m,7m].

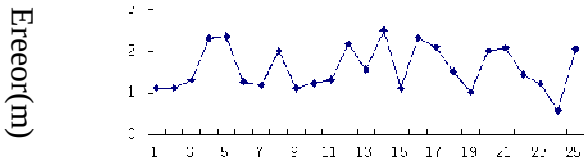

Fig.10 Estimation error of the testing nodes

Conclusions

This paper proposes an improved indoor localization algorithm based on Log-distance path loss model. The two parameters, A and n, of the Log-distance path loss model are estimated by the measured RSSI data in the lab with the length 9.5m and width 6m. Then the two step filtering was carried out to keep the RSSI signals with the higher probability. Finally the Taylor's series iteration method is used to compute the position coordinate of the moving objects. The experiment results show that the location error is no more than 3m.

Acknowledgements

This work was financially supported by special funds for scientific research projects of Marine public welfare industry: (201405022-2).

Reference

[1] Jing Li, Liang Pei, Maoyong Cao, Daoyin Yu, Super-resolution Time Delay Estimation Algorithm Based on the Frequency Domain Channel Model in OFDM SystemsWCICA06，12（6），5144-5148，2006.6.

[2] Jing Li, Liang Pei, Maoyong Cao, Nongliang Sun，Super-resolution Time of Arrival Estimation for Indoor Geolocation based on IEEE 802.11 a/g， WCICA08.

[3] P. Loschmidt, G. Gaderer, and T. Sauter. Clock synchronization for wireless positioning of cots mobile nodes. In IEEE International Symposium on Precision Clock Synchronization for Measurement, Control and Communication, pages 64-69, 2007.

[4] M. Mah, N. Gupta, and A. Agrawala. Pinpoint time difference of arrival for unsynchronized 802.11 wireless cards. In INFOCOM, 2010 Proceedings IEEE, pages 1-5, 2010.

[5] Fang, S. H., T. N. Lin, et al. (2008). Location fingerprinting in a Decorrelated Space, IEEE Transactions on Knowledge and Data Engineering, vol.20, pp.685-691.

[6] Chan, C. L., G. Baciu, et al. (2010). Orientation-based Wi-Fi Positioning on the Google Nexus One, accepted by the 6th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications. Niagara Falls, Canada..

[7] Chan, C. L., G. Baciu, et al. (2009f). Using Wi-Fi Signal Strength to Localize in Wireless Sensor Networks, IEEE International Conference on Communications and Mobile Computing, vol.1, pp.538-542.

[8] Solahuddin Y F, Mardeni R. Indoor empirical path loss prediction model for 2.4 GHz 802.11n network[C]//Control System, Computing and Engineering (ICCSCE),2011 IEEE International Conference on. IEEE, 2011:12-17.

[9] Lott M, Forkel I. A multi-wall-and-floor model for indoor radio propagation[C]//Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd. IEEE, 2001,1:464-468.

[10] Phaiboon S. An empirically based path loss model for indoor wireless channels in laboratory building[C]//TENCON,02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. IEEE, 2002,2:1020-1023. building[C]//TENCON,02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. IEEE, 2002,2:1020-1023.