Indoor Positioning Method Based on Wireless Signal

https://doi.org/10.3991/ijoe.v14i10.9303

Jingjing Yang, Zhenyu Feng, Xuchao Ma, Xiao Zhang^(⊠) Hebei North University, Zhangjiakou, China r78z@foxmail.com

Abstract—In view of the problems of traditional wireless indoor positioning technologies such as errors and a low positioning accuracy that cannot reach the application level required by hospital indoor positioning, this study proposes a hospital indoor positioning method based on wireless signals. This study firstly analyzes the principles of hospital indoor positioning, verifies the reliability and accuracy of the collected data using Gaussian distribution, P-P plot and Q-Q plot, and finally analyzes the collected data using the least square fitting algorithm to obtain a fitting wave attenuation model, which is then applied to the indoor positioning system. Experiments show that this method can reduce the error of indoor positioning in hospitals, and improve the repeatability and measurement accuracy of indoor positioning in hospitals.

Keywords—Indoor positioning, Wireless signal, Gaussian distribution, Fitting wave attenuation model

1 Introduction

The application of various positioning technologies in hospitals has been at the experimental and trial level, and they have gradually been adopted until in recent years. However, because various positioning has its own advantages and disadvantages, only by selecting the appropriate cost-effective solutions and integrating applications with each other, can we obtain more effective precise positioning points and improve practicality.

At present, there are mainly three kinds of indoor positioning principles [1-3] and Table 1 shows the comparison of the different technical principles. At present, in the indoor positioning researches, the mainstream indoor positioning method are mainly based on wireless signal and geometric operation, supplemented by fingerprint operation. At present, the most commonly used indoor positioning technologies include pseudolite, infrared, computer vision, ZigBee [4-6], etc. Except for wireless signal positioning technology, other indoor positioning technologies have disadvantages such as uncontrollable cost, difficulty in development, and mismatched input and output in small locations. Glaces Web [7] et al. have deployed a centralized wireless Zig Bee network for indoor positioning, but the network size is small. Yang [8] et al. uses ibeacon to achieve hospital indoor positioning, but the cost is too high.

Technologies	Principles	Characteristics	Technological path
Geometric positioning	Calculate the current position of the target by measuring the distance between the tar- get and several fixed base stations	Real time opera- tion and high ac- curacy	Base station positioning, wifi positioning, Bluetooth posi- tioning and UWB position- ing, ultrasonic positioning
fingerprint positioning	Pre-collect a physical feature value in dif- ferent locations in the room, draw a fin- gerprint map, locate the feature value of the directly measured location and com- pare it with a fingerprint map	Pre-collection and timely up- date	Wifi fingerprint positioning and geomagnetic positioning
image posi- tioning	Carry out SLAM modeling of indoor scenes by laser radar or camera, and then use machine vision matching algorithm to estimate the location	Specific scene and large amount of oper- ation	laser SLAM and machine vi- sion

Tabla 1	Comparison	ofindoor	nositioning	nringinlag
I abit 1.	Companson	01 ma001	positioning	principies

Due to the popularization and low development costs of wireless hotspots, poor accuracy and high volatility of conventional wireless signal base stations, Gaussian fitting and least square fitting algorithm are used to analyze the collected data to obtain the fitting wave attenuation model as the principle of hospital indoor positioning, which can reduce costs, and improve repeatability and measurement accuracy.

2 The styles

According to the relationship between signal intensity and distance, the traditional wave propagation model is studied [9]. In an ideal space environment, wall blocking and multipath propagation effects are not considered. Assuming that the distance between the wireless signal transmitting end and the wireless signal receiving end is d, the receiving power of the receiving end may be denoted as [10]:

$$P_R(d) = \frac{P_T G_T G_R \lambda^2}{(4\pi)^2 d^2 L}$$
(1)

In the formula (1), P_T is the transmitter transmit power, G_t and G_r are the power amplifications of transmitting and receiving antenna respectively, λ represents the wireless wave wavelength, the unit of P_T and P_T is W, and G_t and G_t are dimensionless. From the above formula, it can be obtained that the distance d^2 is inversely proportional to the receiving power under ideal conditions.

In real environment, there are multi-path propagation, obstacle reflection, signal diffraction and so on, so the ideal wireless propagation loss model is not applicable. Therefore, the logarithmic-normal distribution model is adopted **Fehler! Verweisquelle konnte nicht gefunden werden.**:

$$P_{\gamma}(d) = RSSI = 10 \lg \left[\frac{P_{\gamma}(d_o)}{P}\right] + 20 \lg \left[\frac{d_o}{d}\right]$$
(2)

In the formula (2), d_0 represents the unit distance of 1m, and the unit of P_r is dBm. Considering the complexity of the environment and the positioning accuracy

requirements, the wireless signal attenuation model based on the RSSI ranging algorithm is simplified **Fehler! Verweisquelle konnte nicht gefunden werden.** as:

RSSI=A-10nlgd

(3)

Where, in the formula (3),d is the distance between the wireless base station and the user, whose unit is m. A is the RSSI measured when the distance d between the wireless base station and the user is equal to 1 m, and n is the signal attenuation factor. As different materials have different signal absorption rates, n has different values. Considering the absorption of signals by housing materials, n ranges from 3 to 4.

Since, conventional wave propagation attenuation model has many shortcomings [8], data is collected according to the traditional wave model. According to the principle of statistics, the wave propagation attenuation model is fitted by the least square fitting polynomial based on the traditional model, so as to make it more conform to the actual model.

Under normal circumstances, the calculated received signal strength always has an error from the actual situation, because the field strength is inversely proportional to the distance and the complex terrain of the actual environment leads to weakening signals [11]. In addition to the cause of equipment failure, the distance d from the wireless base station to the user location obtained by the signal attenuation model is generally greater than the real distance. As shown by the left figure in Figure 1, the distances between every two points of user location D and the wireless base stations A, B, and C are calculated respectively based on the RSSI model. A circle is made using the sites and calculated distance, then the intersection point can be obtained.

The basic idea of the centroid positioning algorithm [12] is as follows: Calculate the intersection points of three circles, in which each two are intersecting, then six intersection points are obtained, wherein the three points closest to the user are feature points, which are represented by E, F and G. The feature point is taken as the vertex of the triangle, the centroid of the triangle as the predicted point M where the user is located. The feature points E, F and G are as shown in Figure 1. The calculation method of the feature point E is:

$$\begin{cases} \sqrt{\left(x_{e} - x_{a}\right)^{2} + \left(x_{e} - x_{a}\right)^{2}} \leq r_{a} \\ \sqrt{\left(x_{e} - x_{b}\right)^{2} + \left(x_{e} - x_{a}\right)^{2}} = r_{b} \\ \sqrt{\left(x_{e} - x_{c}\right)^{2} + \left(x_{e} - x_{a}\right)^{2}} = r_{c} \end{cases}$$
(4)

Similarly, the coordinates of F and G can be calculated. At this time, the coordinates of the unknown point are $\left[\frac{x_e + x_f + x_g}{3}, \frac{y_e + y_f + y_g}{3}\right]$.

In Figure 1, D represents the actual location of the user, and M represents the user's location calculated by the centroid algorithm. Then the algorithm is compared with another positioning algorithm, trilateration method, through which the user's coordinate point is estimated to be N. Through the comparison with the two methods, it can be known that the accuracy of the centroid algorithm is higher.



Fig. 1. Trilateration and centroid positioning

3 Fitting Process of Wave Propagation Attenuation Model

The fitting process of the wave propagation attenuation model is mainly divided into two parts, data acquisition and model fitting.

3.1 Data acquisition

The data acquisition part of this design mainly consists of wireless signal base station, mobile phone based on iOS operating system, mac computer based on 10.13 operating system and SQL database. The overall data acquisition structure is shown in Figure 2.

During the measurement, the signal base stations are arranged in a symmetrical manner. In this design, 10 base stations are arranged in a circular manner. The signal acquisition end is in the center of the ring, thus the vertical field intensity vector is cancelled. Set the abscissa of the distance, and the coordinates are from 0.05 to 2.1. 0.05 is the step size of 0.05-0.1, 0.1 is the step size of 0.1-1, and 0.2 is the step size of 0.1-2.1. There of follows: total 17 coordinate points are а as [0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.2, 1.4, 1.6, 1.8, 2, 2.1].



Fig. 2. Overall data acquisition structure

Set the acquisition time interval to 1s, and each RSSI intensity value is collected. When the number is greater than 500 signal points, the distance is moved according to the above table. Data exported from the SQL database is saved as a data file of dat.csv and the actual number of collected data is 89,328.

3.2 Data exploration

Data exploration [13] verifies the reliability and accuracy of this set of data in terms of Gaussian distribution, P-P plots, and Q-Q plots. Gaussian distribution is used to intuitively judge whether Gaussian distribution is met or not, and P-P plots and Q-Q plots are used to verify the correctness of intuition.

Since there are many non-negligible influencing factors in the indoor propagation process of signals, such as multipath propagation effect, background noise, reflection, etc., the relationship between RSSI and the distance D is not a one-to-one correspondence in actual situations, but a one-to-many relationship. It can be seen from the above that the RSSI may dither. Therefore, the reliability of the data can be judged by determining whether the data conforms to the normal distribution. It can be seen from Figure 3 that the data distribution conforms to the Gaussian distribution model.

The P-P plot reflects the relationship between the cumulative ratio of the specified distribution and the cumulative ratio according to the variable. P-P plot can show whether the data conforms to the specified distribution. When all the points in the P-P plot approximately is in a straight line, it indicates that the data conforms to the specified distribution, otherwise it does not. The data of dat.csv is used here.

In statistics, the Q-Q gplot (Q represents the quantile) is a probability plot that compares the two probability distributions graphically to see if they are similar. If the two distributions have linear correlations, most of their points on the Q-Q plot will fall on a straight line. If the studied distributions are in the position-scale category, the parameters can be visually evaluated using a Q-Q plot. First, select the quantile interval, the theoretical quantile interval is -20, and the empirical quantile interval is -10. The dat.csv data is used here.



Fig. 3. Data distribution

It can be judged whether the collected sample data conforms to the normal distribution by checking whether the points on the Q-Q plot fall on a straight line or approximately a straight line. Moreover, the slope of the line reflects the standard deviation, and the intercept is the mean value of the sample. As can be seen from the figure, a small amount of data is slightly shifted at both ends, and a large amount of data in the middle corresponds to a Gaussian distribution. The purpose of Q-Q plot is the same as that of P-P plot, whose difference lies in test methods.

The Q-Q plot and the P-P plot verify the normality of the data, such as the Q-Q plot shown in Figure 4 and the P-P plot shown in Figure 5.



Q-Q plot

Fig. 5. P-P plot

3.3 Data exploration

Since the collected data is discrete, the data is described before data fitting using Box-plot, which is commonly used in statistics.

The preprocessed data is used to draw the Box plot, as shown in Figure 6, where the range, median, and outliers of the preprocessed data can be seen.

Data processing is performed according to the pre-data box-plot. The data in the range of [0.3, 0.7] of each group are retained as shown in Figure 7, and the outliers are removed from the processed data, so that the correspondence of that the data reflect is more clear.



Fig. 6. Box-plot of preprocessed data



Fig. 7. Processed Box-plot

The Box-plot reflects the correspondence between D (Dist) and Rssi to a certain extent.

The result derived from Formula (1) is a theoretical model, which has certain differences from reality. The Box-plot intuitively reflects an approximately linear relationship. However, a logarithmic relationship between Rssi and lgD can be obtained from the theoretical model. So polynomial fitting of lgD and Rssi is carried out. As shown in Figure 8, the green line is a linear fitting of Rssi and D, the yellow line is a linear fitting of Rssi and lgD, the blue line and the red line are the high order fittings of Rssi and lgD. The specific Formulas and their corresponding fitting analysis plots, including Figure 9-1, 9-2, 9-3, and 9-4 are as follows.

The Residuals section provides the main statistics of the prediction error.

The Coefficients section provides the specific values for each parameter and the predictability of each feature. Three stars indicates that this feature cannot be a variable that has nothing to do with the dependent variable, that is, the relationship between the feature and the dependent variable is extremely strong. Estimate is the estimation of the parameter, Std.Error is the standard deviation of the regression parameter sd (parameters), and t value and Pr (>|t|) are all test hypotheses for the regression parameter. t is the t-value of the test hypothesis for the parameter. Pr is used to compare with the significance level to decide whether to accept the hypothesis test.



Fig. 8. Fitting line



fit_4 <- lm(rssi ~ log(dist) + I(log(dist)^2) + I(log(dist)^3), data = dat summary(fit_4) > fit_3 <- lm(rssi ~ log(dist) + I(log(dist)^2), data = dat)
> summary(fit_3) Call: lm(formula = rssi ~ log(dist) + I(log(dist)^2) + I(log(dist)^3), data = dat) Call: lm(formula = rssi ~ log(dist) + I(log(dist)^2), data = dat) Residuals: Residuals: Min 1Q Median 3Q Max -4.4210 -1.3300 0.2586 1.1917 4.1427 Min 1Q Median 3Q Max -4.385 -1.339 0.234 1.216 4.097 Coefficients: Coefficients: LoeTILGENTS: Estimate Std. Error t value Pr(>|t|) (Intercept) -54.427228 0.011604 -4690.5 <2e-16 *** log(dist) -9.290632 0.021029 -441.8 <2e-16 *** I(log(dist)^2) -0.947378 0.009085 -104.3 <2e-16 ***
 Coefficients:
 Estimate Std. Error
 t value Pr(>|t|)

 (Intercept)
 -54.45579
 0.01479
 -3682.463
 2 e-16

 log(dist)
 -9.7064
 0.02198
 -421.716
 2 e-16

 (Iog(dist)^2)
 -0.8485
 0.0301
 -25.706
 2 e-16

 (Log(dist)^2)
 -0.8485
 0.0301
 -25.706
 2 e-16

 (Log(dist)^3)
 0.03141
 0.01008
 3.115
 0.00184
 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ignif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' Residual standard error: 1.73 on 27755 degrees of freedom (61569 observations deleted due to missingness) Multiple R-squared: 0.9468, Adjusted R-squared: 0.9468 F-statistic: 2.47e+05 on 2 and 27755 DF, p-value: < 2.2e-16 Residual standard error: 1.729 on 27754 degrees of freedom (61569 observations deleted due to missingness) Multiple R-squared: 0.9468, Adjusted R-squared: 0.9466 F-statistic: 1.647e+05 on 3 and 27754 DF, p-value: < 2.2e-9-3 9-4

Fig. 9. Formula fitting analysis diagram

The multivariate R-squared (decision coefficient R-squared) provides a measure of model performance **Fehler! Verweisquelle konnte nicht gefunden werden.**, showing that the model interprets the credibility of the change in the dependent variable value from the overall data. The higher the reliability that the fitted model interprets data, the closer the multivariate R-squared value is to 1.0. For example, the multivariate R-squared value of the func-3 function is 0.9468, so 95% of the variation rules of the dependent variable can be explained by the model.

Table 2 summarizes the formula residuals in the above figure. Table 3, Table 4, Table 5, and Table 6 summarize the coefficients of the formulas (5).

	Min	1Q	Median	2Q	Max
5-1	-4.5955	-2.0660	-0.5364	1.4636	10.5715
5-2	-6.4659	-1.0783	0.1791	1.5723	3.0621
5-3	-4.385	-1.339	0.234	1.216	4.097
5-4	-4.4210	-1.3300	0.2586	1.1917	4.1427

Table 2. Residual range of each formula

Table 3. The parameters of Formula 5-1

	Estimate	Std.Error	T value	Pr(> t)
Intercept item	-40.9931	0.0332	-1234.9	<2e-16
Dist	-11.5682	0.0304	-380.6	<2e-16

Table 4. The parameters of Formula 5-2

	Estimate	Std.Error	T value	Pr(> t)
Intercept item	-54.70131	0.01333	-4102.6	<2e-16
Lg(dist)	-7.39960	0.01256	-589.2	<2e-16

Table 5. The parameters of Formula 5-3

	Estimate	Std.Error	T value	Pr(> t)
Intercept item	-54.427228	0.011604	-4690.5	<2e-16
Lg(dist)	-9.290632	0.021029	-441.8	<2e-16
Lg(dist)^2	-0.947378	0.009085	-104.3	<2e-16

Table 6. The parameters of Formula 5-4

	Estimate	Std.Error	T value	Pr(> t)
Intercept item	-54.45579	0.01479	-3682.463	<2e-16
Lg(dist)	-9.27064	0.02198	-421.716	<2e-16
Lg(dist)^2	-0.84852	0.03301	-25.705	<2e-16
Lg(dist)^3	0.03141	0.01008	3.115	0.00184

4 Experiment and Verification

The indoor positioning performance of this design is tested in detail. The macOS High Sierra version 10.13.4 is used on the PC side and the iOS version 11.2.5 is used on the terminal iPhone. In this study, a set of experimental data are analyzed, the results of trilateration centroid positioning algorithm and verification observation algorithm implemented by C code are compared with the actual results, and the errors of the two algorithms are compared. The plan of the experimental site is as shown in Figure 10. The red dot indicates the location of the base station, and the real pictures of the experimental site are shown in Figures 11 and 12, and the space coordinate system is drawn by the red line. The green dotted line represents the space vector for each wireless base station. At (368, 990, 127), data is collected per second for a period of 100 s. Table 7 shows the results of calculation variance and mean square error of partial experimental data and all data.



Fig. 10. Plan of experimental site

Paper-Indoor Positioning Method Based on Wireless Signal



Fig. 11. Real picture of the experimental site



Fig. 12. Real picture of the experimental site

	(x,y)	(X,Y)	Error(cm)	Vari- ance	Mean square devi- ation
	(368,990)	(364.09762,1017.0585)	27.33		
(x, y) as a theoretical	(368,990)	(366.79742,1015.2148)	25.24		
coordinate point	(368,990)	(366.39859,1016.9239)	26.97		
(X, Y) for calculating coordinatesThe lower right cor- ner of a plane map is the origin of the coor- dinates	(368,990)	(364.45340,1014.9879)	25.23	16.29	4.028
	(368,990)	(366.27629,1011.4548)	21.52		
	(368,990)	(367.69000,1017.9368)	27.93	10.58	4.028
	(368,990)	(362.10084,1016.6079)	27.25	-	
	(368,990)	(372.73718,1009.0199)	19.60		
	(368,990)	(365.19362,1019.0593)	29.19		
	(368,990)	(369.77267,1006.5948)	16.68		

Table 7.	Partial	data	and	error	analysis
----------	---------	------	-----	-------	----------

The error data is analyzed. The calculated user coordinate points (x, y) and the user's actual coordinate points (X, Y) are drawn. It can be clearly seen from the Figures 13 that there is not great difference between the calculated user's coordinate points and the actual coordinate points.



Fig. 13. Line chart of positioning error analysis

Conclusions: Through the comparison of the error and variance of the data, it can be seen that the error fluctuation is better than the range of $0.5m \sim 1m$ obtained in the past researches, as the error range of this system is $0.2 \sim 0.7$.

5 Conclusions

As the hospital has strengthened its efforts in informatization and intelligent construction in recent years, the indoor positioning technology studied in this paper will play a more active role in the medical care of the hospital.

6 Acknowledgment

This project was supported partially by Population Health Informatization in Hebei Province Engineering Technology Research Center, Medical Informatics in Hebei Universities Application Technology Research and Development Center and Zhangjiakou Municipal Science and technology and Seismological Bureau (No.1711038C).

7 References

- He, S., Chan, S.H.G. (2016). Wi-Fi fingerprint-based indoor positioning: Recent ad-vances and comparisons. IEEE Communications Surveys & Tutorials, 18(1): 466-490. DOI: 10.1109/COMST.2015.2464084 <u>https://doi.org/10.1109/COMST.2015.2464084</u>
- [2] Cannistraro, M., Bernardo, E. (2017). Monitoring of the indoor microclimate in hospi-tal environments a case study the Papardo hospital in Messina. International Journal of Heat and Technology, 35(1), S456-S465.. <u>https://doi.org/10.18280/ijht.35Sp0162</u>
- [3] Liuzzi, S., Stefanizzi, P. (2016). Experimental study on hygrothermal performances of indoor covering materials. International Journal of Heat and Technology, 34(2), S365-S370. DOI: 10.18280/ijht.34Sp0225. https://doi.org/10.18280/ijht.34Sp0225
- [4] Kanaan, M., Chahine, K. (2018). CFD study of ventilation for indoor multi-zone transformer substation. International Journal of Heat and Technology, 36(1), 88-94. DOI: 10.18280/ijht.360112. <u>https://doi.org/10.18280/ijht.360112</u>
- [5] Yang, C., Shao, H.R. (2015). WiFi-based indoor positioning. IEEE Communications Magazine, 53(3): 150-157. <u>https://doi.org/10.1109/MCOM.2015.7060497</u>
- [6] Hossain, A.K.M.M., Soh, W.S. (2015). A survey of calibration-free indoor positioning systems. Computer Communications, 66: 1-13. <u>https://doi.org/10.1016/j.comcom.2015.03.001</u>
- [7] Yang, H., Yang, S. (2010). Connectionless Indoor Inventory Tracking in Zig Bee RFID Sensor Network. 2009 25th annual conference of IEEE industrial electronic, 1-6: 2470-2475.
- [8] Yang, J., Wang, Z., Zhang X. (2015). An ibeacon-based indoor positioning systems for hospitals. International Journal of Smart Home, 9(7): 161-168. <u>https://doi.org/10.14257/ijsh.20</u> <u>15.9.7.16</u>
- [9] Elson, J., Girod, L., Estrin, D. (2002). Fine-grained network time synchronization us-ing reference broadcasts. ACM SIGOPS Operating Systems Review, 36(SI): 147-163.
- [10] Liu. C., Wu, K., He, T. (2005). Sensor localization with ring overlapping based on comparison of received signal strength indicator. Mobile Ad-hoc and Sensor Systems IEEE International Conference on, 2(56): 516-518. <u>https://doi.org/10.1109/MAHSS.2004.1392193</u>
- [11] Al-Madani B.M., Shahra E.Q. (2018). An Energy Aware Plateform for IoT Indoor Tracking Based on RTPS. Procedia computer science, 130(C): 188-195. <u>https://doi.org/10.1016/j.pro cs.2018.04.029</u>
- [12] Casetti E., Can A. (1986). Trigonometric expansions: An application to the study of migration fields. Modeling and Simulation, 17(1): 23-28.
- [13] Zuur leno E.N., Elphick C.S.A.F. (2010). A protocol for data exploration to avoid common statistical problems. Methods in ecology and evolution, 1(1): 3-14. Little R.J.A. (1988). A test of missing completely at random for multivariate data with miss-ing values. Journal of the American Statistical Association, 83(404): 1198-1202.

8 Authors

Jingjing Yang received his Master Degree from Lanzhou Jiaotong University in 2012. Now he is a doctoral student at the University of Chinese Academy of Sciences and he is currently a lecturer of the school of information science and engineering at Hebei North University. His research interests are in the field of image processing and internet of things.

Zhenyu Feng is the member of Population Health Informatization in Hebei Province Engineering Technology Research Center. His research interests are in the field of internet of things.

Xuchao Ma is the member of Population Health Informatization in Hebei Province Engineering Technology Research Center and Medical Informatics in Hebei Universities Application Technology Research and Development Center. His research interests are in the field of internet of things.

Xiao Zhang is a Professor of the school of information science and engineering, Hebei North University, Zhangjiekou, China. He has published in various National, International Journals and Conference papers. He has 30 years of Teaching Experience. His areas of interests are media security and signal processing.

Article submitted 13 August 2018. Resubmitted 24 September 2018. Final acceptance 23 October 2018. Final version published as submitted by the authors.