Combination of Statistical Access Point Selection Methods Based on RSSI in Indoor Positioning System

Shigeyuki Tateno*, Tong Li*, Yu Wu*, and Ziyuan Wang*

Abstract: Recently, as wireless infrastructure has developed widely, and smartphones have become necessary in daily life, indoor positioning devices and applications have become more and more popular. Previous studies have proposed several methods based on different wireless communication technologies. Among them, methods with received signal strength indicator (RSSI) values and trilateration methods are mainly used to obtain positioning results. However, due to abnormal RSSI values caused by noise and influence from the environment, the accuracies of these methods are not satisfying. Therefore, a new method which can reduce the positioning error is necessary. In this paper, to improve the positioning accuracy above trilateration results, an access point selection method and a kernel density estimation method are combined to obtain estimated points. Experiments are designed in actual environments, and the results of which show that the proposed method is sufficient for improving the positioning accuracy.

Key Words: received signal strength indicator, access point selection, kernel density estimation.

1. Introduction

As wireless fundamental infrastructure is developing widely, and smartphones are becoming necessary in daily life, the using of devices and applications for indoor positioning is growing significantly such as navigation in an underground station, monitoring in healthcare facilities, and so on. There already exist several positioning technologies [1]–[6] like a global positioning system (GPS), an IC tag, a radio frequency identification (RFID) tag, and a beacon. Among them, GPS mainly works outside, and the other devices support very short communication range so that application scenarios are limited to the small area. Besides, according to one of the previous studies, Wi-Fi can estimate long distance between transmitters and receivers, which are based on collecting a received signal strength indicator (RSSI) [7].

The RSSI is a kind of typical distance measurement technique, which represents the power of received radio signal from a transmitting device. The distance between a transmitter and a receiver can be estimated by using the collected an RSSI value and propagation loss. After applying the distance, which calculated from access points (AP) to an unknown node through range-based positioning algorithm, such as trilateration [8],[9], the position of the unknown node can be determined.

The principle of the trilateration is that RSSI values received from access points are converted into distance values, after that the position of the unknown node can be calculated through distances from three different access points. In the conventional trilateration, only the access points with the top three strongest RSSI values are used to conduct the position estimation. However, due to the influence of the environment, the random reflection and refraction of radio wave propagation, RSSI values may become noisy. When these access points with noise are chosen, positioning results will also become noisy.

In this paper, to improve the positioning accuracy above trilateration results, an access point selection method is applied. By adding this method, more access points can be utilized, and the effect of abnormal results can be weakened through the functioning statistical rules. Furthermore, the centroid method and the kernel density estimation (KDE) method are used for analysis of combination results. Experiments are designed in two actual environments to investigate the positioning accuracies of the proposed methods.

The process of the proposed method in this paper has been partially described in the prior conference publication [10]. New access point selection methods are added to the present paper. To evaluate the proposed method, a new experiment in a larger area is also added. Moreover, the influences of the two parameters of the log-distance path loss model used in the prior conference publication are also investigated.

2. System Introduction

The proposed methods are tested on a remote monitoring system. In the system, the RSSI values are gathered from ZigBee devices. In software part, firstly, RSSI data is calculated to obtain distances based on the log-distance path loss model. Then, a weighted trilateration method is used to calculate the estimated point according to information from three access points. After that, statistical AP selection methods, which include the centroid method and the KDE method, are utilized based on the all-combination and the strongest-combination to obtain results from a series of estimated points.

2.1 Hardware

DigiMesh is a kind of mesh protocol based on ZigBee, which is developed by Digi International, Inc. (Digi) [11]. It works on 2.4 GHz frequency band, and the maximum data rate is 250 kbps. The greatest feature of DigiMesh is that there is no hierarchy in the network and all nodes are equivalent.
A DigiMesh network can be built by only routers, while the network cannot be constructed without a coordinator in normal ZigBee network. Moreover, the routers of DigiMesh can also be set into temporary sleep to reduce power consumption, which is not allowed in normal ZigBee network. These characteristics of DigiMesh make it suitable for long-term monitoring.

2.2 Weighted Trilateration

An RSSI represents the strength of received radio signal in dB (mW). The relationship between an RSSI and a distance could be described by the log-distance path loss model [12],[13] as shown in Eq. (1), and Eq. (2) is the transposition of Eq. (1):

\[
R = A - 10N\log_{10}d, \quad (1)
\]
\[
d = 10^{\frac{R}{N}}, \quad (2)
\]

where the received RSSI data is recorded as \(R\), and \(d\) denotes the distance between the target and the access point, \(N\) represents the path loss exponent that reflects the environment, and \(A\) represents the path loss at the reference distance, which is usually set as 1 m.

Trilateration is one of the most widely used algorithms in positioning [14]–[16] since it can be easily realized by existing wireless infrastructures. The position of the target is determined by the intersection point formed by three circles, the radial distance of which is calculated from the RSSI by applying it to Eq. (2) between the target and each access point.

However, the signal propagation in an indoor environment sometimes suffers critical interference. Since the radius of circles are calculated from a signal with noise, there may be more than one intersection points of the three circles. In a real situation, it is probably that each intersection point is only formed by two circles, as Fig. 1 shows.

For the circles without single intersection point, for example, Eq. (3) is established to find the point \(P_t\) in Fig. 1:

\[
(x_1 - x_A)^2 + (y_1 - y_A)^2 \leq r_A^2,
\]
\[
(x_1 - x_B)^2 + (y_1 - y_B)^2 \leq r_B^2,
\]
\[
(x_1 - x_C)^2 + (y_1 - y_C)^2 \leq r_C^2,
\]

where the centers of the three circles \(O_A, O_B, \) and \(O_C,\) are \(A(x_A, y_A), B(x_B, y_B),\) and \(C(x_C, y_C),\) respectively. The distances between the target point \(P_t(x_1, y_1)\) and \(A, B, \) and \(C\) are calculated by Eq. (2) and recorded as \(d_A, d_B, \) and \(d_C,\) respectively. The intersection point \(P_t(x_1, y_1)\) of \(O_A\) and \(O_B,\) is calculated by Eq. (3). Similarly, the intersections \(P_2(x_2, y_2)\) and \(P_3(x_3, y_3)\) can be obtained by the same way as that of \(P_t.\) For the situation of circles without any intersection point, \(P_1, P_2,\) and \(P_3\) are calculated through using a method proposed in [17]. Finally, \(P_T(x_t, y_t)\) is calculated as Eq. (4):

\[
P_T(x_t, y_t) = \frac{P_1(x_A, y_A) + P_2(x_B, y_B) + P_3(x_C, y_C)}{3}.
\]

Since the distance between each access point and the target is estimated based on the RSSI, which means the smaller the distance is, the less the signal loss is, and the higher the accuracy is. Therefore, access points that have different signal strength will affect the distance estimation at different levels which could make a great difference in positioning accuracy. Applying a weighted function as an adjustment is considered to be an effective way to assist the trilateration method. By using a weighted factor \(\omega,\) shown in Eq. (5), Eq. (4) can be adjusted to Eq. (6):

\[
\omega_1 = \frac{1}{r_B + r_C}, \quad \omega_2 = \frac{1}{r_C + r_A}, \quad \omega_3 = \frac{1}{r_A + r_B},
\]
\[
T(x_t, y_t) = \frac{\omega_1 P_1(x_A, y_A) + \omega_2 P_2(x_B, y_B) + \omega_3 P_3(x_C, y_C)}{\omega_1 + \omega_2 + \omega_3}.
\]

2.3 Access Point Selection

A method of exhaustion means that all the possible situations would be listed and tried to solve a specific problem. It can be considered as a simple searching method: to traverse whole elements in a status set in which a feasible solution maybe exists, and to judge whether it is a feasible state. In many cases, a number of situations can be calculated by permutation and combination, which makes it possible to evaluate the difficulty of the problem.

In the experiments, there are three kinds of combination of the access points, for the weighted trilateration generation, are used. Firstly, the all-combination is utilized to list all the possible combinations for the trilateration estimation. Secondly, the strongest-combination is utilized to select the combinations that contain the access point with the maximum RSSI value. For example, when the access point 1 (AP1) has the maximum RSSI strength will a

2.4 Centroid Method

A centroid is the arithmetic mean position of all the points and can be used in calculations such as the geometric center, the center of an area, the centroid of a finite set of points, and so on.

For a finite set of points \(P_1(x_1, y_1), P_2(x_2, y_2), \ldots, P_n(x_n, y_n),\) the centroid point \(P(x, y)\) can be calculated as Eq. (7) shows:

\[
P(x, y) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n}\right).
\]
2.5 Kernel Density Estimation Method

The kernel density estimation (KDE), also called the Parzen-Rosenblatt window method named after Emanuel Parzen and Murray Rosenblatt [18],[19], is a non-parametric method to estimate the probability density function of random variables. It works on a kernel function, which is also used in support vector machine (SVM), and a selected bandwidth. By choosing the suitable kernel function and a bandwidth, a smooth and continuous graph can be drawn in comparison with a histogram.

Consider a univariate independent and identically distributed sample \((x_1, x_2, \ldots, x_n)\) with an unknown density \(f\). The kernel density estimator \(\hat{f}_h\) is shown as Eq. (8):

\[
\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^{n} K_h \left( \frac{x - x_i}{h} \right),
\]

where \(K\) is the kernel function, \(h\) is the bandwidth, and \(K_h\) is the scaled kernel. To yield meaningful estimates, the kernel function should satisfy the following [19]:

\[
\begin{align*}
& \sup_{-\infty < u < \infty} |K(u)| < \infty, \\
& \int_{-\infty}^{\infty} |K(u)| du = 1, \\
& \lim_{u \to \infty} |uK(u)| = 0.
\end{align*}
\]

The most commonly used kernel function is the Gaussian kernel function:

\[
K(u) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} u^2 \right).
\]

KDE can be used in different kinds of access point selection methods to obtain the density distribution. The density distribution results are shown as Fig. 3. The peaks in Fig. 3 indicate the point with the highest possibility of density distribution, which is considered the positioning result in this research.

3. Experiment and Results

The proposed method has been experimentally investigated in a real-life corridor with corners. The experiment has been designed twice in a corridor, one of which is designed in a limited range, and the other is designed in a larger range. The positioning accuracy of the proposed method is also compared with those of the weighted trilateration method and the unilateralation method.

The system used for the experiment firstly gathers the RSSI values from DigiMesh devices and after that calculates the RSSI values for distances based on the log-distance path loss model. Through the weighted trilateration method, the estimated points are generated. Finally, statistical AP selection methods, which include the centroid method and the KDE method, are utilized based on the all-combination and the strongest-combination to obtain results from a series of estimated points.

3.1 Experiment of the Parameter

In the previous research [10], the parameter \(A\) of the log-distance path loss model is set as \(-42\) dB (mW), and the parameter \(N\) of that is set as 2.3. Different parameters values around the previous ones are tried in the proposed method to see whether the positioning results change significantly.

The influence of different \(N\) under constant \(A\), which is \(-42\) dB (mW), is shown in Fig. 4. While \(N\) is larger than 2.2, the results are stable. The results show that different path loss exponents may not infect the positioning results.

The influence of different \(A\) under constant \(N\), which is 2.3, is shown in Fig. 5. While \(A\) is smaller than \(-44\) dB (mW), the
results are stable. The results show that parameter $A$ can be measured previously. Moreover, a suitable and constant value can be used for a real-time experiment.

### 3.2 Limited Range Experiment

Firstly, the proposed method is used in an experiment in a limit range in the corridor shown in Fig. 6, in which 10 units of access points were deployed.

During the experiment, the subject held a DigiMesh device in hand, walked along the route beginning from the start point, and stopped after the first corner, where the total length of the route is 44 m. The route is divided by every 2 m, into totally 23 steps. The map of the route in this experiment is also shown in Fig. 6.

#### 3.2.1 Centroid and kernel density estimation

The centroid method and the KDE method are both used to obtain the positioning points according to all the estimated points with the different combination situations. For the centroid method, the mean coordinates of $x$-axis and $y$-axis are calculated respectively as positioning results. For the KDE method, a Gaussian kernel function is used to estimate the highest probability of density.

#### 3.2.2 Statistical access point selection methods

In this experiment, 10 access points are set, for statistical AP selection methods, the all-combination, the strongest-combination, and the top-6-combination are used. For the all-combination method, each time three access points are chosen to unit one combination. As a result, total 120 combinations are used in this method. The number of combinations can be calculated as Eq. (11) shows.

$$C(10, 3) = C_{10}^3 = 120.$$  \hspace{1cm} (11)

In the case that the target does not receive all the RSSI data from all ten access points, all the available access points the data of which can be received will be used.

For the strongest-combination method, each time the access point with the strongest RSSI is selected to unit one combination with other points. Thus, the number of combinations with the strongest access point is 36, which is shown in Eq. (12).

$$C(9, 2) = C_9^2 = 36.$$  \hspace{1cm} (12)

For the top-$n$-combination method, each time the access points with the $n$th strongest RSSI are selected to unit one combination. Thus, the number of combinations with the top-9, top-8, top-7, and top-6 access points are 84, 56, 35, and 20, respectively.

#### 3.2.3 Result of the experiment

The results of positioning experiment under the centroid method and the KDE method with the all-combination, the strongest-combination, and the top-6-combination, are shown as traces in Fig. 7. Moreover, the results of the centroid method under the all-combination, the strongest-combination, and the top-$n$-combination, where $n$ ranges from 9 to 5, are shown in Table 1. The results of the KDE method under the same combinations are shown in Table 2. These tables show the minimum, maximum, average, and standard deviation of the positioning errors.

From the results, it can be concluded that the KDE method has smaller error than the centroid method. On the other hand, the strongest-combination is better than the all-combination while the best result appears at the top-6-combination.

### 3.3 Larger Range Experiment

Secondly, the proposed method is used in an experiment in a large range in the corridor shown in Fig. 8, in which 27 units of access points were deployed.

![Fig. 5 Influence of parameter $A$ on location error under constant parameter $N$.](image1)

![Fig. 6 Experiment environment and route in limited range experiment.](image2)

![Fig. 7 The positioning tracing results of the limited range experiment.](image3)
access points were deployed.

During the experiment, the subject held a DigiMesh device in hand, walked along the route beginning from the start point, and stopped after the third corner; the total length of the route is 150 m. The route is divided by every 2 m, into totally 76 steps. The map of the route in this experiment is also shown in Fig. 8.

In this experiment, for the reason that some steps can only obtain RSSI values from less than eight access points, it is hard to calculate KDE under the strongest-combination. Thus, only the top-7, the top-6, the top-5 and the all-combinations are compared.

The results of positioning experiment under the centroid method and the KDE method with the all-combination and the top-6-combination, are shown as traces in Fig. 9. Moreover, the results of the centroid method under the all-combination and the top-\(n\)-combination, where \(n\) ranges from 7 to 5, are shown in Table 3. The results of the KDE method under the same combinations are shown in Table 4.

From the results, it can be concluded that the KDE method has smaller error than the centroid method. On the other hand, the best result still appears at the top-6-combination in terms of the minimum and the average of the positioning error.

| Combination  | Min (m) | Max (m) | Ave (m) | SD (m) |
|--------------|---------|---------|---------|--------|
| All          | 1.20    | 9.78    | 3.97    | 1.99   |
| Strongest    | 0.38    | 7.56    | 2.85    | 1.47   |
| Top-9        | 0.26    | 7.46    | 3.50    | 1.82   |
| Top-8        | 0.40    | 7.17    | 3.14    | 1.62   |
| Top-7        | 0.48    | 5.27    | 2.83    | 1.35   |
| Top-6        | 0.04    | 5.76    | 2.31    | 1.46   |
| Top-5        | 0.48    | 5.94    | 2.45    | 1.42   |

Table 3 Errors of centroid methods in larger range experiment.

| Combination  | Min (m) | Max (m) | Ave (m) | SD (m) |
|--------------|---------|---------|---------|--------|
| All          | 1.96    | 19.65   | 9.14    | 3.51   |
| Top-7        | 0.61    | 11.67   | 5.28    | 2.39   |
| Top-6        | 0.80    | 9.92    | 4.69    | 2.30   |
| Top-5        | 0.38    | 9.84    | 4.41    | 2.44   |

Table 4 Errors of KDE methods in larger range experiment.
3.4 Comparison of the Proposed Methods

The positioning accuracy of the proposed method is also compared with the weighted trilateration method and the unilateration method.

The result of the weighted trilateration method and the unilateration method in the limited range experiment and the larger range experiment are shown as traces in Figs. 10 and 11. Moreover, the results of the comparisons with the KDE method under the proposed top-6-combination are shown in Tables 5 and 6.

From the results, it can be concluded that the KDE method is better in terms of the maximum and the standard deviation of the positioning error, which means it has higher stability.

4. Conclusion

In this research, the influence of parameters in the log-distance path loss model is tested. To improve the positioning result of the trilateration, access point selection methods have been proposed. Moreover, to calculate the positioning result through the selected access points, the centroid method and the kernel density estimation method are used. The experiment for the parameters in the model shows that there is no obvious change in positioning accuracy when the parameter $A$ is below $-44$ dB (mW) and the parameter $N$ is above 2.2, which means these parameters can be set as constant values. Experiments designed in different environments show that the positioning results of the kernel density estimation method have higher accuracy than those of the centroid method, while the positioning results of the strongest-combination have higher accuracy than those of the all-combination. The highest accuracy in this research appears in the top-6-combination. Comparison results of the trilateration method, the unilateration method, and the kernel density estimation method show that, although not always having the best average positioning error, the kernel density estimation method under the proposed top-6-combination is better in terms of the maximum and the standard deviation of the positioning error, which means it has higher stability. For future work, since the accuracy can be improved by the kernel density estimation method, other statistical methods should be considered. Moreover, experiments in more complex environments still need to be considered.

References

[1] Y. Shigematsu, S. Aikawa, and Y. Nakamura: An empirical study on the compose method of outdoor indoor position estimation, *IEEE Transactions on Electronics, Information and Systems*, Vol. 138, No. 2, pp. 140–148, 2018.

[2] S. Nirjon, J. Liu, G. DeJean, B. Priyantha, Y. Jin, and T. Hart: COIN-GPS: Indoor localization from direct GPS receiving, *Proc. International Conference on Mobile Systems, Applications, and Services* 2014, pp. 301–314, 2014.

[3] M. Oehliai, H. Hatano, M. Fujii, and A. Itou: Indoor positioning based on fingerprinting method by incoming GPS signals, *IEICE Trans. Fundamentals*, Vol. E99-A, No. 1, pp. 319–322, 2016.

[4] T. Fuse and K. Matsumoto: Development of self-localization method using sensors on mobile devices, *Journal of the Japan Society of Photogrammetry and Remote Sensing*, Vol. 54, No. 6, pp. 290–299, 2016.

[5] S.S. Chawathe: Beacon placement for indoor localization using Bluetooth, *Proc. International IEEE Conference on Intelligent Transportation Systems* 2008, pp. 980–985, 2008.

[6] H. Seki, H. Sakashita, Y. Kamiya, M. Hikizu, and H. Nomura: On the technology of acoustic utility. Acoustic information and energy. Navigation system for an indoor mobile robot by active ultrasonic beacons, *Transactions of the Japan Society of Mechanical Engineers Series C*, Vol. 67, No. 657, pp. 1282–1288, 2001.

[7] Y. Zhu, W. Xia, Y. Zhang, F. Yan, X. Zuo, and L. Shen: A hybrid indoor localization algorithm based on RSSI and inertial navigation, *Telecommunications Science*, Vol. 2017, No. 10, pp. 99–106, 2017.

[8] L. Zhou and S. Ma: Improved hybrid localization algorithm of maximum likelihood and centroid localization based on RSSI, *Proc. International Conference on Information Science and Control Engineering*, pp. 385–389, 2017.

[9] Y. Ge, Z. Zheng, B. Yan, J. Yang, Y. Yang, and H. Meng: An RSSI-based localization method with outlier suppress for wireless sensor networks, *Proc. International Computer and Communications* 2016, pp. 2235–2239, 2016.

[10] S. Tateno, T. Li, Y. Wu, and Z. Wang: Improved indoor localization system using statistical AP selection method based on RSSI, *Proc. SICE Annual Conference* 2018, pp. 1652–1657, 2018.

[11] Digi International Inc., XBee/XBee-PRO S2C DigiMesh 2.4 Radio Frequency (RF) Modules, 90001506 datasheet, Aug. 2016 [Revised June 2017].
timal access points selection method, Proc. SICE Annual Conference 2017, pp. 321–326, 2017.

[13] W. Xue, W. Qiu, X. Hua, and K. Yu: Improved Wi-Fi RSSI measurement for indoor localization, IEEE Sensors Journal, Vol. 17, No. 7, pp. 2224–2230, 2017.

[14] A.S. Paul and E.A. Wan: RSSI-based indoor localization and tracking using sigma-point Kalman smoother, IEEE Journal of Selected Topics in Signal Processing, Vol. 3, No. 5, pp. 860–873, 2009.

[15] C. Pei, Y. Cai, and Z. Ma: An indoor positioning algorithm based on received signal strength of WLAN, Proc. Pacific-Asia Conference on Circuits, Communications and Systems 2009, pp. 516–519, 2009.

[16] M. Khan, Y. Kai, and H.U. Gul: Indoor Wi-Fi positioning algorithm based on combination of location fingerprint and unscented Kalman filter, Proc. International Bhurban Conference on Applied Sciences and Technology 2017, pp. 693–698, 2017.

[17] J. Li, X. Yue, J. Chen, and F. Deng: A novel robust trilateration method applied to ultra-wide bandwidth location systems, Sensors, Vol. 17, No. 4, 795, 2017.

[18] M. Rosenblatt: Remarks on some nonparametric estimates of a density function, The Annals of Mathematical Statistics, Vol. 27, No. 3, pp. 832–837, 1956.

[19] E. Parzen: On estimation of a probability density function and mode, The Annals of Mathematical Statistics, Vol. 33, No. 3, pp. 1065–1076, 1962.

Shigeyuki Tateno (Member)

He received his B.S., M.S., and Ph.D. degrees from Kyushu University, Japan, in 1990, 1992, and 2001, respectively. In 2003, he joined Waseda University, where he is currently an Associate Professor of Graduate School of Information, Production and Systems. His research interests include wireless network systems.

Tong Li

He received his B.Eng. degree from Xidian University, China, in 2015. In 2017, he joined Waseda University, Japan, where he is currently a student of Graduate School of Information, Production and Systems. His research interests include wireless network systems.

Yu Wu

She received her B.Eng. degree from Beijing Institute of Technology, China, in 2016. In the same year, she joined Waseda University, where she is currently a student of Graduate School of Information, Production and Systems. Her research interests include indoor localization.

Ziyuan Wang

He received his B.Eng. degree from Xidian University, China, in 2017. In the same year, he joined Waseda University, where he is currently a student of Graduate School of Information, Production and Systems. His research interests include computer vision and data analysis.