An Interpolation Method of Indoor Positioning for Modern Breeding Management

Jin Ren1,2*
1School of Information Science and Technology, North China University of Technology, Beijing100144, P.R.China
2Beijing Key Laboratory on Integration and Analysis of Large-scale Stream Data, Beijing100144, P.R.China
*Corresponding author (E-mail: rj@ncut.edu.cn)

Songyang Huang
School of Information Science and Technology, North China University of Technology, Beijing100144, P.R.China

Wei Song
School of Information Engineering, Minzu University of China, Beijing100081, P.R.China

Yunan Wang
School of Information Science and Technology, North China University of Technology, Beijing100144, P.R.China

Fuqiang Song
School of Information Science and Technology, North China University of Technology, Beijing100144, P.R.China

Abstract
Indoor positioning has become a hot research topic for modern breeding management. Wi-Fi is a common public facility in life, which can be used for indoor positioning. The commonly used Knn algorithm has the problem of low computational efficiency. A method is proposed to apply a clustering algorithm to data classification, which can improve efficiency. However, with the increase in the number of classifications, the problem of insufficient data may occur in positioning. In order to avoid too little data after classification and affect the subsequent calculation, this paper proposes a new method to supplement the data by using the interpolation algorithm. The experimental results show that the location accuracy of the optimized algorithm is better than that of the single clustering algorithm.

Key words: Interpolation Method, Indoor Positioning, Clustering Algorithm, Modern Breeding Management.

1. Introduction
In the recent years, with the rapid development of wireless communication technology, Location-based services are more and more widely used in daily life, like modern breeding management. Modern breeding management requires precise location of personnel or livestock. Location technology is a necessary link and core technology to realize location-based services. Due to the attenuation of the signal in the process of propagation, if the outdoor positioning method is applied to the interior, it will be limited. At present, the indoor positioning technology based on Wi-Fi has been widely concerned. With the development of technology, Wi-Fi has covered most homes and public places. Modern breeding management needs to be carried out indoors as well as outdoors. In order to provide Wi-Fi location service, access points (APs) will be deployed indoors. Each wireless AP has a globally unique MAC address. In general, the wireless AP does not change for some time. The Wi-Fi terminal can continuously scan and collect the surrounding AP signals and obtain their signal strength.

Fingerprint positioning is a common method of Wi-Fi indoor positioning. Compared with traditional TDOA (Time Difference of Arrival) and AOA (Angle of Arrival) algorithms, fingerprint positioning has advantages such as no additional hardware equipment and low cost. Wifi based fingerprint positioning is usually divided into two stages: offline collection and online positioning. In the offline stage, $\text{rss}_i$ should be collected from each access point in the location area using an instrument capable of detecting $\text{rss}_i$ signals and the
fingerprint database should be classified and generated. In order to make the collected data stable and reliable, the data are usually typically filtered. In the online stage, the real-time measured data is compared with the fingerprint database, and the final result of localization is obtained by a series of algorithms.

Location-based fingerprinting is usually divided into two types. One of the types is the deterministic algorithm, which compares signal characteristics (such as vector \( r \)) and the calculated statistics in the fingerprint database. The other type is the probabilistic algorithm, which is used to calculate the probability that a signal feature belongs to a certain distribution (stored in a fingerprint database). In this article, we mainly study the deterministic positioning algorithms.

2. Traditional Algorithm

2.1. Knn Algorithm

Knn algorithm is the most commonly used algorithm in fingerprint positioning, which summarizes the coordinates of the nearest \( k \) points and takes the average to get the location \([1, 2]\). The algorithm is shown in equation (1) and (2).

\[
(d_j) = \left(\sum_{i=1}^{n} r_{ss_i} - r_{ss_j}\right)^2 \\
(\bar{x}, \bar{y}) = \frac{1}{k} \sum_{i=1}^{k} (x_i, y_i)
\]

where \( n \) represents the number of access points detected at the target location, \( r_{ss_i} \) represents the signal strength of the test point received by APs, \( r_{ss_j} \) represents the signal strength of the reference point’s fingerprint in the fingerprint database. \( d_j \) represents the Euclidean distance between the test point and the \( j \)-th reference point. \( k \) represents the number of reference points, \( (\bar{x}, \bar{y}) \) represents the location coordinates of the reference point \( i \).

Although the Knn algorithm is simple and effective, it also has some defects [3]. Because of the test point is performed with all the fingerprint points in the database, the calculation efficiency is low.

2.2. Clustering Algorithm

The clustering algorithm is one of the methods to solve the problem of positioning efficiency [4]. Its main purpose is to classify the data according to the characteristics of it [5].

In the fingerprint location, firstly, the fingerprint database is clustered according to the feature of \( r_{ss_i} \), and the clustering center of each class is marked. Then, the appropriate class for positioning is selected by comparing the Euclidean distance between the measured point and the center of convergence. In this way, the clustering algorithm can reduce the number of fingerprint points needed and improve computational efficiency [6].

The commonly used clustering algorithms include K-means algorithm, FCM (Fuzzy C-means) algorithm [7], PAM (Partitioning Around Medoids) algorithm and so on. Different clustering algorithms cluster data sets according to different rules [8].

Although the clustering algorithm can improve the computational efficiency, when there are too many classifications, it will cause the loss of key data and affect the positioning accuracy. In this case, we will analyze and solve this problem.

3. Proposed Method

Due to the shortcomings of the clustering algorithm mentioned earlier, we have taken a new approach to supplement the data. In this paper, a method is proposed to supplement the clustered fingerprint points by two-dimensional interpolation.

After the database is classified, the number of fingerprint points contained in each dataset is not equal. For example, we cluster data points to four categories on a two-dimensional plane. The result is shown in Figure 1. The solid dots in Figure 1 represents the clustered fingerprint points, and the triangle represents the measured point. All fingerprint data are divided into 4 categories. Based on the results of the calculation, category 3 will be selected as the new database. All fingerprint data are divided into 4 categories, and the points of the remaining three categories will not be used. Since this class contains fewer fingerprint points, there will be a large distance error when the KNN algorithm is executed. Therefore, we proposed to interpolate the classified data. The schematic diagram after interpolation is shown in Figure 2. The solid rectangles represent interpolated data points. After interpolation, the fingerprint points in this area are denser than before. The accuracy of positioning will also increase. The general function of two-dimensional interpolation is in equation (3).

\[
z = f(x, y)
\]
where \((x,y)\) are the coordinates of the known fingerprint points, \(z\) is the rss vector corresponding to the coordinate. \(f(\cdot)\) represents the algorithm formula of four two-dimensional interpolation methods\([9, 10, 11]\). The meaning of two-dimensional interpolation is to obtain the unknown rss vectors from the known coordinates.

\[
f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \\ f_4(x) \end{bmatrix}
\]

\[
T \left[ \begin{array}{c} f_1(x) \\ f_2(x) \\ f_3(x) \\ f_4(x) \end{array} \right] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

\[
T_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

\[
T_\beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

\[
T_\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

\[
T_\delta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

where \(T\) represents the transformation matrix.\[T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

\[f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \\ f_4(x) \end{bmatrix} = T \begin{bmatrix} (x_1,y_1) \\ (x_2,y_2) \\ (x_3,y_3) \\ (x_4,y_4) \end{bmatrix}
\]

3.1. Types of Two-dimensional Interpolation

Two-dimensional interpolation methods generally include neighboring point interpolation (Nearest), bilinear interpolation (Bilinear), cubic spline interpolation (Spline) and double cube interpolation (Cube). The following is an introduction to each of them.

1) Nearest Interpolation

The nearest neighbor interpolation is the simplest interpolation method. In two-dimensional space, the coordinates of fingerprint points are known, so the method is to select the nearest fingerprint point to the target point, find the data numerical value of it and assign the same numerical value.

2) Bilinear Interpolation

Bilinear interpolation is one of the most commonly used algorithms in the two-dimensional interpolation model. The core idea is to perform linear interpolation in two directions. The calculation of bilinear interpolation is based on the rss of four fingerprint points around the target point \(o(x,y)\), \(f(x,y)\) is the rss of \(o(x,y)\), which is calculated by twice interpolation algorithm. The schematic diagram is shown in Figure 3.

First of all, according to the positioning coordinate of the target point, it is divided into the corresponding clustering. For example, as shown in Figure 3, the target point \(o(x,y)\) is positioned between a square surrounded by four fingerprint points, Their coordinates are \((m,n),(m+1,n),(m,n+1)\) and \((m+1,n+1)\). The question now is how to calculate the rss at \(o(x,y)\) based on the number of rss at \((m,n),(m+1,n),(m,n+1)\) and \((m+1,n+1)\). The bilinear interpolation technique uses the following method: it is assumed that the change of rss in the space is linear in the longitudinal direction, so that \(f(m,y)\) and \(f(m+1,y)\) at the coordinates of \((m,y)\) and \((m+1,y)\) can be obtained according to the linear equation or geometric proportional relation. Then, suppose that the change of rss is still linear on the straight line determined by \([((m,y),f(m,y)])\) and \([(m+1,y),f(m+1,y)])\). The linear equation is obtained, and then as the calculated numerical value of rss at the target point \(o(x,y)\), \(f(x,y)\) can be obtained. This is the basic idea of bilinear interpolation. The two basic assumptions used are: first, rss is linear in the longitudinal direction, and then it is assumed that the rss is also linear in the transverse direction. The specific process is shown in Table 1.
Table 1. Proposed bilinear interpolation algorithm

<table>
<thead>
<tr>
<th>Interpretation</th>
<th>Equation</th>
</tr>
</thead>
</table>
| After clustering, Select a fingerprint point neighbor to the target point in  | \[
\begin{aligned}
\alpha &= x - m \\
\beta &= y - n \\
\end{aligned}
\]  |
| the corresponding class. Calculate \( \alpha \) and \( \beta \), the distance   | (4)                                                                      |
| between them in horizontal and vertical coordinates.                          |                                                                          |
| Calculate \( f(m, y) \) based on \( f(m, n) \) and \( f(m, n+1) \)           | \[
f(m, y) = f(m, n) + \alpha [f(m, n) - f(m, n+1)]
\]  |
| Then calculate \( f(m+1, y) \) based on \( f(m+1, n) \) and \( f(m+1, n+1) \) | \[
f(m+1, y) = f(m+1, n) + \alpha [f(m+1, n) - f(m+1, n+1)]
\]  |
| interpolation.                                                                | (5)                                                                      |
| Finally calculate \( f(x, y) \) by twice interpolation based on \( f(m, y) \) | \[
f(x, y) = f(x, y) + \beta [f(m+1, y) - f(m, y)]
\]  |
| and \( f(m + 1, y) \).                                                        | (6)                                                                      |
|                  |                                                                 |
| Combine formula (4)(5) with (6), get the result.                              | \[
f(x, y) = f(x, y) + \beta [f(m+1, y) - f(m, y)]
\]  |
|                  |                                                                 |
|                  |                                                                 |
|                  |                                                                 |
|                  |                                                                 |
| Therefore, the rss\(_i\) of any target point after clustering can be obtained  |                                                                          |
| according to formula (8) base on the rss\(_i\) of the four fingerprint points   |                                                                          |
| around it.                                                                   |                                                                          |

![Figure 3. Schematic diagram of bilinear interpolation algorithm](image-url)

3) **Cubic Spline Interpolation**

Cubic spline interpolation is calculating through cubic polynomials. The calculation process of cubic spline interpolation is more complex, the specific process is as follows Table 2.

Table 2. Proposed cubic spline interpolation algorithm

<table>
<thead>
<tr>
<th>Interpretation</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>First, calculate the one-dimensional case. Set up function ( S(x) ).</td>
<td>( x \in [a, b] )</td>
</tr>
<tr>
<td>( S(x) ) is a cubic polynomial on each interval, ( x_0, x_1, \ldots, x_n )</td>
<td>( a = x_0 &lt; x_1 &lt; \cdots &lt; x_n = b )</td>
</tr>
<tr>
<td>are the horizontal ordinates of known fingerprint points.</td>
<td></td>
</tr>
<tr>
<td>( z_j ) are their rss(_i) values. ( S(x) ) is the cubic spline function</td>
<td>( z_j = S(x_j), j = 0, 1, \ldots, n. )</td>
</tr>
<tr>
<td>of the known fingerprint points.</td>
<td></td>
</tr>
</tbody>
</table>
Assume the second derivative of cubic Spline interpolation function is \( S''(x) \).

\[
S''(x) = M_j, \quad (j = 0, 1, \ldots, n) \quad (12)
\]

Assume the distance of interval is \( h_j \).

\[
h_j = x_{j+1} - x_j, \quad (j = 0, 1, \ldots, n - 1). \quad (13)
\]

Because of \( S(x) \) is a cubic polynomial on interval \([x_j, x_{j+1}]\). So \( S''(x) \) is a linear function in the interval \([x_j, x_{j+1}]\), which can be expressed as the follow equation (14).

\[
S''(x) = M_j \frac{x_{j+1} - x}{h_j} + M_{j+1} \frac{x - x_j}{h_j} \quad (14)
\]

Integrate \( S''(x) \) twice, calculate the integral constant base on \( S(x_j) = z_j \) and \( S(x_{j+1}) = z_{j+1} \). Then get the expression of cubic spline interpolation function.

\[
S(x) = M_j \frac{(x_{j+1} - x)^3}{6h_j} + M_{j+1} \frac{(x - x_j)^3}{6h_j} \left(z_j - \frac{M_j h_j^2}{6} x_{j+1} - x \right) + \frac{M_{j+1} h_j^2}{6} \frac{x - x_j}{h_j} \quad (15)
\]

The \( M_j \) above is unknown. In order to determine \( M_j \), the first derivative of \( S(x) \) needs to be calculated.

\[
S'(x_j + 0) = S'(x_j - 0) \quad (16)
\]

Calculate \( M_j \),

\[
\mu_j M_{j-2} + 2M_j + \lambda_j M_{j+1} = d_j \quad (17)
\]

Calculate \( \mu_j, \lambda_j \) and \( d_j \) above.

\[
\mu_j = \frac{h_{j-1}}{h_{j-1} + h_j}, \quad \lambda_j = \frac{h_j}{h_{j-1} + h_j}, \quad d_j = \int \frac{f[x_j, x_{j+1}] - f[x_{j-1}, x_j]}{h_{j-1} + h_j} \quad (18)
\]

---

**Figure 4.** Positioning flow chart of the proposed method

Because of \( S(x) \) is a cubic polynomial, four undetermined coefficients need to be determined on each interval \([x_j, x_{j+1}]\), and there are \( n \) intervals, therefore, \( 4n \) parameters should be determined. According to the continuity of the second derivative of \( S(x) \) in \([a, b] \), \( 3n-3 \) conditions can be provided if the continuity conditions of the first and second derivatives at point \( x \) are locally satisfied. In addition, \( S(x) \) satisfies the interpolation condition, there are \( 4n-2 \) conditions, so two additional conditions are needed in order to determine \( S(x) \). A boundary condition can be added to the endpoint \( a = x_0, b = x_n \) of the interval \([a, b] \). For example, the first
derivative at both ends can be used to provide two additional conditions. \( M_j \) is obtained by the combining of above \( 4n \) equations, and the interpolation function \( S(x) \) can be determined. Through two times interpolation, one dimension is selected at a time, the other dimension is maintained, and the selected dimension is calculated by one-dimensional interpolation. Finally, after twice interpolation, the corresponding interpolation result can be calculated based on the interpolation function \( S(x) \).

4) Double Cube Interpolation

Double cube interpolation is the most commonly used interpolation method in two-dimensional space. The required interpolation data is obtained by calculating the weighted average of the nearest sixteen sample points in a rectangular grid.

The results of the four interpolation algorithms are different, we will analyze the interpolation effect in the experiment. The overall positioning process is shown in Figure 4. We will simulate and analyze the algorithm according to this process.

3. Experiment

We obtain positioning fingerprint data in feeding farm. The test area is about 20 meters long and 15 meters wide. There are 6 access points in this area. Divide the area into 300 square cells and receive the test \( rss_i \) in the lower right corner of each cell, so we obtained 266 \((19*14)\) distributed reference points in this space. Each fingerprint point contains a vector consisting of 6 \( rss \) vectors. We obtained fingerprint data. We selected 30 fingerprint point as test points, the rest of fingerprint points as reference points. The average positioning error obtained from multiple tests is used as the performance index.

Under the same conditions, we compare the positioning effects of the traditional Knn algorithm, the FCM clustering algorithm and the improved double cube interpolation algorithm. We choose the number of classification as 4, the number of interpolation points is twice as many as the number of fingerprints after classification. Then the statistics and analysis of the error results obtained by each positioning algorithm are carried out. The conclusion is drawn through the comparison of the image and the table.

![Empirical CDF](image)

**Figure 5.** Probability density comparison of three positioning algorithms

**Table 3. Comparison of three positioning algorithms**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Rmse (m)</th>
<th>With in 3m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knn</td>
<td>2.12</td>
<td>80%</td>
</tr>
<tr>
<td>FCM</td>
<td>2.01</td>
<td>83%</td>
</tr>
<tr>
<td>Interp+FCM</td>
<td>1.72</td>
<td>86%</td>
</tr>
</tbody>
</table>

Firstly, we analyze the improvement effect of the interpolation algorithm (Interp) and the FCM algorithm selected by the clustering algorithm. The error probability distribution of the positioning algorithm is shown in Figure 5. From the Figure 5 we can see that the combination of clustering algorithm and interpolation algorithm has better positioning effect than traditional clustering algorithm. Table 3 provides a statistical analysis of the data of the three algorithms. The average error of the improved interpolation algorithm is 0.31m lower than that.
of the FCM clustering algorithm. In terms of a probability distribution, 86% of the fingerprint point positioning error of the improved interpolation algorithm is less than 3m, which is 3% higher than the FCM algorithm.

Then we compare the adaptability of the three algorithms to noise. Figure 6 shows the performance comparison of three algorithms after adding different intensities of Gaussian noises. From the figure, we can see that the improved interpolation algorithm has better adaptability to noise than the other two algorithms. As the intensity of the noise increases, the error increase of the interpolation improvement algorithm is relatively flat.

\[ \text{Rmse} = \sum (x_i - \hat{x}_i)^2 

In addition, we also compare four different methods in the two-dimensional interpolation algorithm. The comparison effect is shown in Figure 7 and Table 4. Data analysis shows that linear interpolation is the best method in this positioning under the same conditions.

For nearest interpolation, find the distance of the closest fingerprint point from the database and assign the same value to the target point, because this method is too simple, the final calculation effect is not ideal. The result of bilinear interpolation is much better than that of nearest neighbor interpolation, because the nearest neighbor interpolation only makes use of the coordinate of one fingerprint point around, while the bilinear interpolation makes use of the coordinates of four points around it.

Because the boundary conditions need to be determined in spline interpolation calculation, the selection of boundary conditions is involved. A suitable boundary condition will greatly improve the accuracy of the calculation. In practice, because there are only fingerprint points used for interpolation and the specific function is unknown, as a result, it is difficult to select the most appropriate boundary, resulting in an increase in the error. For double cube interpolation, which is based on the average of the nearest sixteen fingerprint points,
because of the clustering process, there will be a large error if the target point is not at the center of the category, but near the boundary. So it is not the more reference points, the better.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Rmse (m)</th>
<th>With in 3m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
<td>1.72</td>
<td>86%</td>
</tr>
<tr>
<td>Bilinear</td>
<td>1.70</td>
<td>90%</td>
</tr>
<tr>
<td>Nearest</td>
<td>1.78</td>
<td>80%</td>
</tr>
<tr>
<td>Spline</td>
<td>1.75</td>
<td>84%</td>
</tr>
</tbody>
</table>

4. Conclusions

In order to solve the problem of insufficient data after classification by the clustering algorithm, this experiment uses a two-dimensional interpolation algorithm to improve and fill the data. The experimental results show that the proposed method is effective and feasible, and the accuracy of positioning is improved by 0.31m compared with the single clustering algorithm. This article provides a new solution for the indoor positioning technology of modern breeding management.

Acknowledgements

This work was supported by 2013 technology foundation for selected overseas chinese scholar (No.401053740756), ministry of personnel of beijing and the scientific research foundation for the returned overseas chinese scholars, state education ministry and training program for Outstanding Young Scholars (No.14085), 2017 annual “outstanding young teacher training program” project of north china university of technology (No.XN019009), scientific research project of beijing educational committee (No.KM201710009004), 2017 youth scientific research innovation fund project of north china university of technology (No.1743024). 2018 science and technology activities project for college students of north china university of technology. Research project on teaching reform and curriculum construction of North China University of technology (No. NCUT2018JGM07). 2019 beijing university student scientific research and entrepreneurship action plan project.

References