Adaptive Density Graph-Based Manifold Alignment for Fingerprinting Indoor Localization

SHIBAO LI, QINGSHUAI ZHAO, HE LI, SHENGZHI WANG, TINGPEI HUANG, AND JIANHANG LIU
College of Oceanography and Space Informatics, China University of Petroleum, Qingdao 266555, China

ABSTRACT The received signal strength (RSS) fingerprint-based indoor localization has been considered as a promising solution, due to its relatively high localization accuracy and its ease of use in widespread Wireless Local Area Network (WLAN) infrastructure. A major bottleneck is that the offline fingerprint calibration is time consuming and labor intensive. In this study, inspired by our analysis that multi-density is inherent to the RSS distribution, we present a new radio map construction scheme, called Adaptive Density Graph-based Manifold Alignment (ADG-MA), which can reduce the number of Reference Points (RPs) in offline phase. In particular, it utilizes the density features to capture the exact neighborhood relations of RSS. Furthermore, the approach labels the RSS from user traces to construct the radio map. The extensive experiments demonstrate that the proposed method can construct an accurate radio map at a low deployment cost, as well as achieve a high localization accuracy.

INDEX TERMS Fingerprint indoor localization, radio map construction, weighted graph construction, manifold alignment.

I. INTRODUCTION

With the unceasing development of wireless communication technology, computer technology and the demand of Location-based Services (LBSs), the indoor wireless localization techniques are of growing interest and becoming increasingly prevalent [1]. Indoor localization has a wide range of applications in indoor navigation and localization, industrial goods management, fixed-point market advertising delivery, game development and smart home. So far, many techniques have been applied to indoor localization, such as infrared ray, ultrasound, radio frequency identification device (RFID) [2], Bluetooth, ultra-wideband (UWB), magnetic field and Wireless Local Area Network (WLAN) [3]. At present, the localization technology based on WLAN mainly includes localization based on Time of Arrival (TOA) [4], Time Difference of Arrival (TDOA) [5], Angle of Arrival (AOA) [6] and Received Signal Strength (RSS) [7]. Compared with other localization algorithms, WLAN fingerprint-based localization scheme is widely used because of its advantages such as low transmission bandwidth requirements, high localization accuracy and no need for additional hardware deployment.

The WLAN fingerprint-based localization approach mainly involves two phases, namely offline phase and online phase [8]. In offline phase, several Reference Points (RPs) are selected uniformly or randomly in the survey area. Then, a survey is conducted and multiple copies of RSS measurements are read at each RP from available Access Points (APs) to construct the radio map. Due to the complexity of wireless Wi-Fi signal propagation and the effect of multi-path fading and shadow effect, the RSS signal of the same AP at a certain RP usually has complex time-changing characteristics, so it needs to collect data repeatedly at the RP to improve the accuracy of the radio map. In the online phase, the newly-collected RSS signals are matched against the signals in the radio map by matching algorithms (e.g. KNN [9]) to calculate the location of the user.

It is obvious that it needs to collect and train numerous labeled RSS data in the offline process. For example, in the process of constructing radio map of a 3.6 m × 7.2 m small office area using the traditional radio map construction
The main contribution of this paper is as follows. In Section II, we review the previous related work. In Section III, the characteristics of multi-density distribution of RSS in the localization scenes are analyzed. In Section IV, the algorithm of constructing adaptive weighted graph for multi-density data is proposed, and it applies to constructing radio map with manifold alignment. In Section V, the analysis based on experiments is used to verify the performance of the proposed radio map construction approach. In Section VI, we give the conclusion of this paper.

II. RELATED WORK

A. THE CONSTRUCTION APPROACHES FOR RADIO MAP

Generally, the radio map stores the mathematical characteristics of RSS collected at the certain RP. The deterministic approaches usually store the mean value, maximum value, etc. Probabilistic algorithms usually store the statistical characteristics of the RSS at RP, and they are mainly divided into two types, parametric and non-parametric estimation. For instance, the radio map in [11] stores histogram of whole data. The radio map in [12] stores the Gaussian distribution parameters. Because the indoor environment is complex and varied, in order to improve the accuracy of localization, many researchers have put forward different forms of radio map based on empirical statistics. The authors in [13] propose that the RSS distribution is non-Gaussian and left-skewed. The radio map stores the corresponding parameters of the distribution characteristics. However, no matter what features are stored in the radio map, in the offline phase, it requires to collect numerous RSS data from plenty of RPs, which needs a huge manual cost.

For this disadvantage of fingerprint-based localization, the researchers proposed propagation model-based construction algorithms [14], dynamic radio map construction [15], crowdsourcing approaches [16]–[18] and the construction methods based on semi-supervised manifold learning [19]–[22]. The crowdsourcing approaches use many non-professional users to contribute locations and RSS data information, and build a radio map through a server. However, due to the introduction of non-professional data, this method introduces a large number of abnormal data. In addition, the system also needs to filter the malicious users interference. It may increase the processing burden of the system and reduce the accuracy of the radio map.
B. THE SEMI-SUPERVISED LEARNING APPROACHES FOR LOCALIZATION

The semi-supervised learning method achieves the calibration of a large number of unlabeled data through a limited number of labeled RSS. During the construction of the radio map, the unlabeled RSS data can be easily collected through random traces creating by users. Based on this advantage, the authors in [19] use semi-supervised learning to construct the radio map. The idea is to obtain the coordinates of the unlabeled data by the positions of a limited number of labeled RSS data, and to realize the construction of the hybrid radio map. The semi-supervised manifold learning method not only utilizes the distance information represented by the RSS value to locate the user, but also considers the manifold geometry features of the RSS data. At the same time, the prior information of the labeled RSS data can be fully utilized. And the manifold learning can capture the nonlinear features of RSS for localization. Therefore, due to these potential characteristics of manifold learning methods and their advantages in localization accuracy, they have received a lot of attention in recent years.

The authors in [20] propose a method of using the manifold regularization to obtain the positions of mobile nodes and APs through joint collaborative filtering, while the experiment is only based on the Line-of-sight (LOS) condition. The authors in [21] use the manifold learning to introduce coordinate dataset and RSS dataset. Through the correspondence between the data in two datasets, the corresponding low-dimensional eigenvectors are calculated, and the position of unlabeled data can be obtained through the similarity of eigenvectors. However, the coordinate dataset in this method needs to accurately reflect the indoor layout structure. In general, the semi-supervised manifold alignment method considers Laplacian graph as a discrete approximation of RSS manifold structure [23]. The authors in [22] propose a hybrid semi-supervised manifold alignment radio map construction method, and meanwhile utilize RSS data to construct Laplacian graph. It preserves the similarities of high-dimensional RSS by the timestamp information of unlabeled data, and completes construction of the radio map through manifold alignment.

However, Laplacian graph used in [20], [22] usually utilizes the k-Nearest Neighbor graph (k-NN graph) to build the weight matrix without considering the effect of data density. Due to the nearest neighbor number k, it is a global fixed factor. When the data is sparse, the data vertices in the weighted graph are over-connected [24]. In this situation, the distance between the neighbor points is quite large and the weighted graph are over-connected [24]. In this situation, the distance between the neighbor points is quite large and it cannot accurately reflect the neighbor relationship of the data points. It is not conducive to the spread of labeled coordinates in the process of manifold alignment, and it affects the performance of indoor localization. Instead of using k-NN graph by conventional methods, the proposed ADG-MA utilizes local density of RSS and shared nearest neighbor similarity to construct the Laplacian graph. Meanwhile, this paper uses area labels of all RSS data to weaken neighbor connections of dissimilar data, ensuring that the weighted graph can accurately reflect the neighborhood relations of all RSS data, which can improve the accuracy of the radio map.

III. THE FEATURE ANALYSIS OF RSS DATA DISTRIBUTION

A. THE FEATURE ANALYSIS

The application scenarios of indoor localization tend to be complicated. In this type of scenarios, there are numerous APs and the wireless communication environment is changeable. The dimension of the RSS vector in the localization process is equal to the number of APs, which indicates that the high dimension RSS data is stored in the radio map. However, the features of the high-dimensional data are complex, and it is difficult to select an appropriate algorithm to extract the RSS features. In this section, the density parameter method and the data dimension reduction algorithm called Linear Discriminant Analysis (LDA) [25] are used to analyze the RSS data in the complex scenarios, and the distribution characteristics of the high dimension RSS data are obtained.

According to the data density parameter proposed by [26], the $\rho_i$ is defined as

$$\rho_i = \sum_j \chi(d_{ij} - d_c)$$  (1)

where $\chi(x) = 1$ if $x < 0$ and $\chi(x) = 0$ otherwise. $d_{ij}$ represents the distances between RSS data points. $d_c$ represents the cutoff distance. The data density $\rho_i$ denotes the number of data that are closer than $d_c$ to the data point $i$. $d_c$ is an empirical parameter, which has good robustness to adapt to large data sets. We utilize $d_c$ to calculate the data density $\rho_i$ which is a standard of the distributions of RSS density about different datasets. We first analyze the TUT dataset [27], UJIindoorLoc dataset [28] and RSS fingerprint data in our experiment scene of an office building. Then, we calculate the distribution of data density and the results are as shown in Fig. 1. The values of inter-quartile range are 25,16,13 respectively, which presents the span of the data density is large. Fig. 2 shows the data density of 120 RPs in TUT dataset. It can be seen that the density parameters of different RPs are quite different.

Further, the distribution characteristics of the projection in 2D of high-dimensional RSS data are analyzed. The RSS data of 2 rooms on the second floor of UJIindoorLoc dataset
are processed by LDA. LDA is a fundamental data analysis approach based on Fisher’s criterion, which is also a supervised method. It can obtain projections of high-dimensional data in the low-dimensional space by minimizing the within-class scatter and the maximum between-class scatter during the dimensionality reduction process. The between-class $S_B$ and within-class $S_W$ scatter matrices are defined as

$$S_B = \frac{1}{N} \sum_{i=1}^{N} (r_i - \bar{r})(r_i - \bar{r})^T$$  \hspace{1cm} (2)$$

$$S_W = \frac{1}{P} \sum_{i=1}^{N} \sum_{m=1}^{n} (r_i(m) - r_i)(r_i(m) - r_i)^T$$  \hspace{1cm} (3)$$

where $N$ represents the number of RPs, $n$ represents the number of APs, $P$ represents the number of all RSS vectors, $r_i$ and $\bar{r}$ represent the mean of RSS at $i$-th fingerprint point and the mean value of all RSS vectors respectively. $r_i(m)$ represents the $m$-th RSS vector of $i$-th RP. Then, it can reduce the dimension of data to a chosen $L = 2$ by projecting onto the linear subspace. The optimal result can be obtained by computing the following objective function:

$$\arg \max_{\Gamma} \left( \frac{\text{tr}(\Gamma^T S_B \Gamma)}{\text{tr}(\Gamma^T S_W \Gamma)} \right)$$  \hspace{1cm} (4)$$

where $\Gamma$ is the basis vector of subspace. Then, we can get the LDA projection of RSS samples in 2D space as shown in Fig. 3. As we can see, the RSS data are divided into two clusters, and the clusters show a significant density difference.

In summary, the RSS data in the localization scene show obvious multi-density distribution characteristics in high dimension space and low dimension projection space. The uneven density of RSS data is common in multiple localization scenes, which affects the accuracy of the localization. This paper aims to design a weighted graph construction algorithm adapted to the multi-density distribution of RSS data to improve the accuracy of labeling the data in the manifold alignment process.

### B. THE MULTI-DENSITY FEATURE

The RSS multi-density feature is used to characterize the non-uniformity of the distribution of RSS vectors in high-dimensional data space. Specifically, in the same distance range, RSS vectors in dense regions have more neighbors. Similarly, in the sparse regions, there is a big difference in the distances between neighbors. The multi-density feature of RSS distribution is shown as Fig. 4. Different points in the figure represent different types of RSS vectors. The thickness of the lines between points indicates the similarity of points in the weight graph. Fig. 4(a) shows three types of RSS points with uneven distribution in the high-dimensional space. The points in the same type are closer than the points in different types, it means that the connections between RSS vectors of same type are stronger than that between RSS vectors of different type. Fig. 4(b) is the result of the weight graph which is created in the conventional method [29]. As the dotted line shows, only considering the distance metric and ignoring the RSS multi-density feature would result in a tighter connection between different types of RSS data, which in turn would affect the accuracy of the weight graph.

In this paper, we propose to use the local RSS density factor, shared nearest neighbor similarity and area information to create the neighbor weight matrix. Then, the manifold alignment is performed for the construction of radio map. The architecture of the proposed system is shown in Fig. 5. In the offline phase, we use a mobile phone to measure RSS from $n$ APs, and the number of RPs in target environment is $N$, the number of RSS collected by multiple
user traces is \( M \). Less APs lead to the decrease of positioning accuracy and more APs result in the increase of positioning complexity. This paper mainly analyzes the performance under reasonable APs distribution. The total collected RSS can be represented as \( R = (r_1, r_2, \ldots, r_{N+M})^T \), where \( r_i = (r_{i1}, r_{i2}, \ldots, r_{in}) \) (\( i \in [1, N+M] \)) represents all RSS data at the \( i \)-th point. \( R^l = (r_1, r_2, \ldots, r_N)^T \) represents the labeled RSS data collected at a certain RP and \( R^u = (r_{N+1}, r_{N+2}, \ldots, r_{M})^T \) represents the unlabeled RSS data of users traces respectively. The coordinates can be represented as \( P = (p_1, p_2, \ldots, p_{N+M})^T \), where \( P^l = (p_1, p_2, \ldots, p_{N})^T \) stands for the coordinates of RPs corresponding to \( R^l \), and \( P^u \) stands for the coordinates of unlabeled data. We initially set the random values to \( P^u \). Firstly, the mixed data matrix \( R \) is made up of the labeled data \( R^l \) and the unlabeled data \( R^u \). According to the multi-density characteristics of the RSS data and the area information, the adaptive local density matrix \( W_d \) is constructed, and the geometric structure model of the RSS manifold is obtained as \( G = (R, W_d) \). Secondly, the multi-density graph Laplacian is constructed on the basis of \( W_d \). The locations of unlabeled data are calibrated by manifold alignment, and the radio map is constructed with the whole RSS data. Finally, the user location estimation can be realized by KNN.

**B. ADAPTIVE LOCAL DENSITY WEIGHTED GRAPH CONSTRUCTION**

1) LOCAL DENSITY SIMILARITY MATRIX CONSTRUCTION

According to Section III, the RSS data in the localization scene have different local density characteristics. The local density factor is calculated by the distance statistics of the neighbor points, and the local density factor [30] is calculated by

\[
\delta_i = d_s(r_i, r_k) = \frac{\sum_{p=1}^{N} |r_{ip} - r_{kp}|}{\sum_{p=1}^{N} (r_{ip} + r_{kp})}
\]

where \( \delta_i \) represents the Sorensen distance [29] of \( r_i \) to its \( k \)-th nearest neighbor point \( r_k \). Considering the impact of local density, the distance from \( r_i \) to \( r_j \) is defined as \( d_{ij} = d_s(r_i, r_j)/\delta_i \), while the converse is \( d_{ji} = d_s(r_j, r_i)/\delta_j \).

Normally, the square distance is calculated as \( d^2_s(r_i, r_j)/\delta_i \), and the similarity of \( r_i \) and \( r_j \) can be represented by the radial basis function (RBF)

\[
A_{ij} = \begin{cases} 
\exp \left( -\frac{d^2_s(r_i, r_j)}{\alpha_1 \delta_i \delta_j} \right), & \text{if } r_i \in N_k(r_j) \text{ or } r_j \in N_k(r_i) \\
0, & \text{otherwise}
\end{cases}
\]

where \( \alpha \) is the user control factor of the RBF. Only when the adjusted \( d_{ij} \) and \( d_{ji} \) are both small, it can obtain the bigger value. When \( d_s \) is smaller than \( \delta_i \) and \( \delta_j \), \( r_i \) is more similar to \( r_j \). Then, (5) can ensure that the similar \( r_i \) and \( r_j \) in the high-dimensional space have the same magnitude of density.

2) SHARED NEAREST NEIGHBOR SIMILARITY

Similar RSS vectors in high-dimensional space have more common neighbors. In order to further prevent excessive connections between dissimilar RSS vectors, the similarity is calculated by

\[
S(r_i, r_j) = |N_k(r_i) \cap N_k(r_j)|
\]

where \( N_k(r_i) \subseteq R \). Substituting the above equation into (6), the similarity of \( r_i \) and \( r_j \) can be defined as

\[
\hat{A}_{ij} = \begin{cases} 
\exp \left( -\frac{d^2_s(r_i, r_j)}{S(r_i, r_j) \delta_i \delta_j} \right), & \text{if } r_i \in N_k(r_j) \text{ or } r_j \in N_k(r_i) \\
0, & \text{otherwise}
\end{cases}
\]

3) THE WEIGHTED GRAPH CONSTRUCTION OF AREA INFORMATION OPTIMIZATION

In the offline phase, different data are collected from different areas. Therefore, area information stands for the potential labeled messages. Although the samples that in the same physical sub-region are not necessary belong to the same cluster after clustered in the signal space. The RSS data in different areas are more likely to have less similarity. We can divide different areas by k-means. We can set the area label of \( r_i \) as area \((r_i)\) and compute weight matrix by

\[
W_{ij} = \begin{cases} 
\theta(r_i, r_j) \exp \left( -\frac{d^2_s(r_i, r_j)}{S(r_i, r_j) \delta_i \delta_j} \right), & \text{if } r_i \in N_k(r_j) \text{ or } r_j \in N_k(r_i) \\
0, & \text{otherwise}
\end{cases}
\]

where \( \theta(r_i, r_j) \) is area weighting factor, which is defined to reduce the connection weight of data from different areas. If \( r_i \) and \( r_j \) have the same area label, \( \theta(r_i, r_j) = \gamma_1 = 1 \); if not, \( \theta(r_i, r_j) = \gamma_2 < 1 \). Finally, we can get adaptive local density neighborhood weighted graph \( G = (R, W_d) \) and weight matrix \( W_d = (W_{ij})_{(N+M) \times (N+M)} \).

Fig. 6 provides the effect of the proposed method. x, y and z indicate three sets of RSS points in two clusters which satisfy
$d_i(x, y) = d_i(x, z)$. We use lines to connect the neighbors of points, x and y have common neighbors which are connected by them. In Fig. 6(a), two clusters of RSS data have the similar densities. $S(r_x, r_y) > S(r_z, r_z)$, which indicates that x and y have more common neighbors. We can get that x and y have larger connecting weight according to (8). In Fig. 6(b), the RSS data have a huge density difference. Fig. 6b shows that $\delta_x < \delta_x$ and $\delta_z \approx \delta_y$. We can get that the weight between x and z is smaller than the weight between x and y according to (8), so the connection between x and z is weaker.

The first part in (10) is the fitting error, which penalizes the difference between the low-dimensional embedding coordinates of RSS and the location coordinates. And the second one is the regularized manifold constraint, which can preserve the structure of the RSS manifold in the low-dimensional space to enforce smoothness along the manifold. $\mu$ is a tradeoff parameter.

Notice that $\frac{\partial C}{\partial dC} = (L_d + L_d^T) C$, the above problems can be solved by standard linear algebra, and we can get the optimal solution by taking the derivative of the right side of (10)

$$C = \left( \mu I_k + L_{ll} \quad L_{lu} \quad L_{uu} \right)^{-1} \left( \mu I_k \quad 0 \right) P$$

$$L_d = \left( \begin{array}{ll} L_{ll} & L_{lu} \\ L_{ul} & L_{uu} \end{array} \right)$$

is the joint graph Laplacian matrix.

$$W_d = \left( \begin{array}{ll} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{array} \right)$$

is the multi-density weight matrix obtained in Section IV-B. The first N-th rows and M-th columns of $W_d$ indicate the neighborhood weight with respect to labeled RSS data. The last M-th rows and M-th columns of $W_d$ indicate the neighbor connection weight with respect to RSS data which is collected through random traces. $W_{lu}$ indicates the neighborhood weight between labeled RSS data and unlabeled RSS data collected through random traces. $I_k = \text{diag}(\phi_1, \phi_2, \ldots, \phi_{N+M})$ is an indication matrix, where $\phi_i = 1$ if the coordinate of the $r_i$ is given and otherwise $\phi_i = 0$. At last, the latter M rows of mapping matrix C are the low dimensional mapping coordinates of unlabeled RSS data.

C. ADG-MA LOCALIZATION

In offline phase, the low-dimensional mapping coordinates of the collected high-dimensional RSS data $\hat{R}$ are assumed to be $C = (c_1, c_2, \ldots, c_{N+M})$. In order to maintain the manifold structure and calculate the low dimensional mapping coordinates of unlabeled RSS data, the Laplacian Eigenmap [31] is introduced. Then we get optimal mapping by:

$$\sum_{i,j} \|e_i - e_j\|^2 W_d = \sum_{i,j} \left( c_i^T c_i - 2c_i^T c_j + c_j^T c_j \right) W_d$$

$$= 2 \sum_{i,j} D_{ij} c_i^T c_i - 2 \sum_{i,j} c_i^T c_j W_d$$

$$= 2 \text{tr} \left( CD_d C^T \right) - 2 \text{tr} \left( CW_d C^T \right)$$

$$= 2 \text{tr} \left( C \left( D_d - W_d \right) C^T \right)$$

$$= 2 \text{tr} \left( C L_d C^T \right)$$

(10)

where $\text{tr} (\bullet)$ presents the trace of the matrix, $L_d = D_d - W_d$ is the Laplacian matrix [32], which is also a positive semidefinite matrix, and $D_d$ is a diagonal matrix with $D_d = \sum_{j=1}^{N+M} W_d (i, j)$.

Considering the RSS manifold, the prior low dimensional manifold coordinates of subset $R'$ are known. It means that pairwise correspondences between $R'$ and $P'$ are known. We construct the following convex differentiable objective function as

$$F(C) = \arg \min_C \sum_{i=1}^{N} \mu \|c_i - p_i\|^2 + C^T L_d C$$

(11)

V. EXPERIMENTAL RESULTS

A. EXPERIMENT SETTING

To evaluate the performance of the proposed algorithm in real environments, the test bed is setup in a building located in China University of Petroleum (East China) as
shown in Fig. 7. The area with the dimensions of 70.2 m × 18 m represents a typical office environment as it includes several labs, offices and one corridor. For our experiments, we divide the target environment into 6 areas. We use an application on the mobile phone (Xiao Mi 5X with Android 7.1) to collect the RSS data from 5 TP-LINK-WR886N APs. For comparison, 360 RSS data (with the sampling rate of 1 Hz) are collected at each RP (with the total number of 412, i.e. ≈ 1.4 fingerprint/m²) one by one from four orientations and 1145 unlabeled RSS data are recorded on user traces.

The RSS collected in each area is described as follows. From Area 1 to Area 6, the number of point is (37,103), (38,110), (95,261), (115,318), (70,192), (57,161), which the first number represents the labeled RSS data and the second one is unlabeled RSS.

**B. LOCALIZATION PERFORMANCE**

In this section, we focus on the localization performance by using the proposed manifold radio map and other four radio maps constructed by RADAR [9], KNN-graph-based Manifold Alignment (KG-MA) [33], Locally Linear Embedding-based Manifold Alignment (LLE-MA) [21] and LE-MA [22].

Fig. 8 depicts the corresponding cumulative distribution function (CDF) of localization errors for each radio maps. When the localization error is 3.5m, the cumulative probability of ADG-MA, KG-MA, LLE-MA, LE-MA and Radar are 74.2%, 57.5%, 57.0%, 66.0% and 52.0% respectively. From this result, we can find that the proposed radio map outperforms the other methods in localization accuracy. As discussed above, the proposed algorithm could capture the potential multi-density feature of RSS data to construct graph Laplacian which reduces the effect of density on the RSS similarity in a complex indoor environment. The conventional manifold alignment approaches only utilize global parameter k to compute the neighborhood weight matrices. Thus, the negative influence of density causes a decline in the localization accuracy.

Fig. 9 compares the CDFs of localization errors by using the proposed algorithm with and without area labels. From this figure, it can be seen that the localization accuracy with area labels is at least 5% higher than the one without area labels. It can be inferred that the area labels contain the inherent neighborhood relations of RSS samples, which is beneficial to calculate the coordinate of unlabeled RSS data.

Fig. 10 depicts the percentage degradation in mean localization errors performance, compared to the reference of localization accuracy with full fingerprints. In comparison with the full fingerprints performance, the results show that our proposed algorithm can achieve a degradation in localization error of 28% as the ratio of labeled RSS is reduced by 80%, while the one by the RADAR, LLE-MA and LE-MA achieve degradations of 103%, 50% and 33% respectively. The ADG-MA algorithm is more robust to the reduction of the ratio of labeled RPs and can ensure the localization accuracy.
C. THE COST OF CONSTRUCT RADIO MAP

Fig. 11 shows the contrast of the manual time cost of the three radio map construction methods with the ratio of the labeled RPs increasing. When the ratio of labeled RPs is 1, in order to capture RSS variation property, the traditional radio map construction requires collecting 360 RSS samples at each one of 412 labeled RPs from four orientations, the time cost for the RADAR radio map construction is 41.2 (\(=412 \times 360/3600\))h. However, the proposed method only requires 20 RSS samples each RP, and the total time cost is 2.6 (\(=(412 \times 20 + 1145)/3600\))h. Meanwhile, the computational time for the manifold algorithm is approximately 42 seconds, and it can be ignored for the reason that it is too small relative to the manual time cost.

D. SELECTION OF PARAMETERS FOR LOCALIZATION

Fig. 12 shows the mean localization errors with respect to \(\gamma^2\). As can be seen from this figure, the mean localization error reaches minimum when \(\gamma^2 = 0.8\), therefore, the weight of area label is selected to be 0.8 in our system.

VI. CONCLUSION

In this paper, a novel approach of reducing the cost of radio map construction has been proposed. We present the novel neighborhood weight matrix to capture the multi-density feature and connection relations of RSS. In addition, we utilize it to obtain the coordinates of the unlabeled RSS to construct the radio map through manifold alignment. And it is confirmed that the proposed method can cut down the deployment cost for the offline radio map construction and guarantee the localization accuracy. The experiments verify the efficiency and effectiveness of the proposed ADG-MA. In future, we will also characterize the local RSS geometries using the other manifold approaches to further improve the precision and stability of radio map.

REFERENCES

[1] R. Melamed, “Indoor localization: Challenges and opportunities,” in Proc. Int. Conf. Mobile Softw. Eng. Syst., May 2016, pp. 1–2.
[2] J. Wang, Y. Ma, Y. Zhao, and K. Liu, “A multipath mitigation localization algorithm based on MDS for passive UHF RFID,” IEEE Commun. Lett., vol. 19, no. 9, pp. 1652–1655, Sep. 2015.
[3] F. Yang, Q. Zhai, G. Chen, A. C. Champion, J. Zhu, and D. Xuan, “Flash–Loc: Flashing mobile phones for accurate indoor localization,” in Proc. IEEE 35th Annu. IEEE Int. Conf. Comput. Commun. (INFOCOM), Apr. 2016, pp. 1–9.
[4] J. He, S. Li, K. Pahlavan, and Q. Wang, “A realtime testbed for performance evaluation of indoor TOA location system,” in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2012, pp. 482–486.
[5] B. Huang, S. Gu, and W. Xia, “A model of cloud computing-based TDOA location system,” in Wireless Communications, Networking and Applications. Springer, 2016, pp. 1187–1194.
[6] A. Fascista, G. Ciccarese, A. Coluccia, and G. Ricci, “A localization algorithm based on V2I communications and AOA estimation,” IEEE Signal Process. Lett., vol. 24, no. 1, pp. 126–130, Jan. 2017.
TINGPEI HUANG received the B.S. and M.S. degrees in computer science from the University of Petroleum, China, in 2004 and 2007, respectively, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2013. She was a Postdoctoral Researcher with the University of Chinese Academy of Sciences, Beijing, from 2013 to 2015. She is currently a Research Assistant with the College of Computer and Communication Engineering, China University of Petroleum. Her main research interests are related to cross layer design and network protocols for wireless networks, especially for wireless sensor networks, 802.11 networks, and the Internet of Things.

JIANHANG LIU was born in 1978. He received Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, in 2013. He visited the University of Maryland, from December 2016 to January 2018. He is currently an Associate Professor with the China University of Petroleum, Qingdao, China. His research interests include the Internet of Things, vehicle networks, and wireless networks.