

RESEARCH

Open Access



# Indoor location method of interference source based on deep learning of spectrum fingerprint features in Smart Cyber-Physical systems

Yunfei Chen\*, Taihang Du\*, Chundong Jiang and Shuguang Sun

## Abstract

The intensity acquisition and fluctuation of the signal intensity of the interference source caused by the indoor multipath effect are very great, and there is a problem that the best eigenvalue is difficult to choose. A kind of unsupervised machine learning algorithm is proposed, which can independently identify and select the optimal eigenvalue without relying on the prior information. First, the wave signal filtering is reduced and processed by kernelized principle component analysis (KPCA) algorithm. Then, the eigenvalues are selected and the redundant features are eliminated by adaptive parameter adjustment denoising auto-encoder (APADAE) algorithm. Finally, the feature vectors are classified and identified by Softmax algorithm and the classification process are optimized by the particle swarm optimization (PSO) algorithm. Experimental results of the Smart Cyber-Physical systems show that the algorithm can indirectly improve the accuracy of the source location based on improving the classification accuracy.

**Keywords:** Indoor positioning, Deep learning, Denoising auto-encoder (DAE), Kernelized principle component analysis (KPCA), Smart Cyber-Physical systems, Spectrometer receiver

## 1 Introduction

Because of the small size, low power, and portability of new communication jammers such as pseudo-base stations, short interference duration, and high randomness of interference area, it is very difficult to supervise them. Interference source is essentially a radio signal transmitter. It covers the frequency band of mobile phone and other communication tools with a large signal intensity. Through the right-signing loophole of a mobile phone, it compels it to link with the pseudo-base station, passively receives spam messages, or leaks security information. People's lives are deeply affected by it, which causes the regulatory authorities to attach great importance to it.

There are few researches on the location of interference sources [1] in indoor environment. At present, fingerprint method [2] is often used to extract fingerprint features of signals in different locations. Indoor radio wave multipath

effect makes the signals fluctuate greatly. How to extract fingerprint features effectively is very important to the location accuracy. Spectrum research technology for communication signal sources is more advanced, but mostly outdoor research, such as satellite, radar, communication base station, and other radiation sources [3] spectrum research. In the field of indoor positioning, fingerprint enhancement is usually achieved by adjusting the access point (AP) mode. The maximum matching method proposed by Zhang et al. [4] aimed to select the best AP combination to improve the positioning accuracy. The premise of the method is that there are sufficient AP resources in the region. Reference [5] applied principal component analysis (PCA) and used AP with the highest contribution rate to locate. On the basis of it, Liu et al. [6] improved Kalman filter method, improved the accuracy of signal acquisition, and thus improved the overall positioning performance. Yin et al. [7] proposed to optimize the positioning efficiency according to AP energy consumption. On this basis, the water injection model AP

\* Correspondence: [chanffett@163.com](mailto:chanffett@163.com); [thdu@hebut.edu.cn](mailto:thdu@hebut.edu.cn)  
School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China

optimized deployment method [8] enhanced fingerprint characteristics to improve the overall positioning efficiency. It deployed up to seven positioning APs in a small positioning environment. The enhancement of fingerprint features in the above literature mainly improves the efficiency of the localization system by increasing the contribution rate of AP matching, but none of them can optimize the system with a small number of AP. In the actual location of the interference source, each additional 1 AP will increase the cost and personnel deployment of the amount of labor collected.

In recent years, artificial intelligence [9–13] and machine learning technology [14, 15] have been popularized and applied, in which depth learning training test model [16] is more significant and has been widely used in image resolution pattern recognition and other fields. The quality of feature selection determines the generalization performance. The contributions of this paper are as follows:

- 1) The spectrum characteristics of interference source signal are taken as the research object in the study.
- 2) A novel unsupervised machine learning algorithm based on traditional multi-layer denoising auto-encoder (DAE) [17] is proposed to extract AP multi-position spectrum features from the localized coordinate points. It can independently distinguish and select the optimal unlabeled data features without prior information.
- 3) In the end, the dimension reduction subset of the labeled data is extracted by KPCA [18]. And we explored deep learning model training, testing, and classifying the output data to improve the positioning accuracy.

The rest of this paper is organized as follows. Section 2 discusses the methods and indoor interference source location system, followed by the selection and extraction of eigenvalues of spectrum signals models in Section 3. The classifier construction, optimization, and training test are discussed in Section 4. Section 5 shows the choice of experimental equipment and the experimental results. Section 6 concludes the paper with summary and future research directions.

## 2 Indoor interference source location system and pretreatment

### 2.1 Methods

This study originates from the need of finding illegal radio stations, and pseudo-base stations have become the focus of the regulatory authorities. Radio interference sources are essentially radio signal transmitters. The shielding effect of radio waves makes it very difficult to locate interference sources in indoor environment. At

present, indoor location is widely applied in location fingerprint information matching method.

In recent years, there have been many studies on fingerprint positioning methods in the field of indoor positioning. The fingerprint positioning method mainly include two stages: off-line location fingerprint acquisition to establish database and online positioning. The main task of the off-line phase is to establish the location fingerprint information database by detecting the signal intensity of the location area with a radio spectrum analyzer. In the online positioning stage, the detected signal intensity value is directly compared with the fingerprint database in the server through the algorithm, and then, the position coordinates of the signal transmitter can be inferred.

### 2.2 Indoor interference source location system and pretreatment

Traditional fingerprint localization methods usually deploy multiple AP in the localization area, sample interference sources at a reference point (RP) location at the same time, and collect all AP received signal strength indicators (RSSI) to the server as fingerprint characteristics of the point coordinates. There are two difficulties in indoor fingerprint localization of interference sources. First, spectrometer receiver is not suitable for large-scale deployment because of its high price and constrained by the field environment. Second, there are many kinds of interference sources with complex spectrum characteristics, and its optimal characteristic parameters are not easy to choose, so it is not suitable to use traditional fingerprint method to characterize the location label.

In view of the above problems, this paper proposes a new feature selection localization method. The localization system includes two stages: off-line acquisition and online localization, as shown in Fig. 1.

In the off-line acquisition stage, the interference source is deployed in a fixed RP position, and the system only calls an AP to detect the spectrum signal. Test data at multiple different AP locations around each RP coordinate. That is to say, the spectrum fingerprint features of different points around RP are extracted to represent an RP coordinate of the interference source. In order to build a positioning system, we choose the learning classification strategy training test.

In the online positioning stage, the random position signal samples of the interference source are collected on the spot, and the input system is compared with the database feature space. The coordinates of the interference source are identified by the classifier according to the RP position class label.

### 2.3 Off-line acquisition of interference source signal spectrum

The spectrum characteristics of interference source signal mainly include transient spectrum characteristics

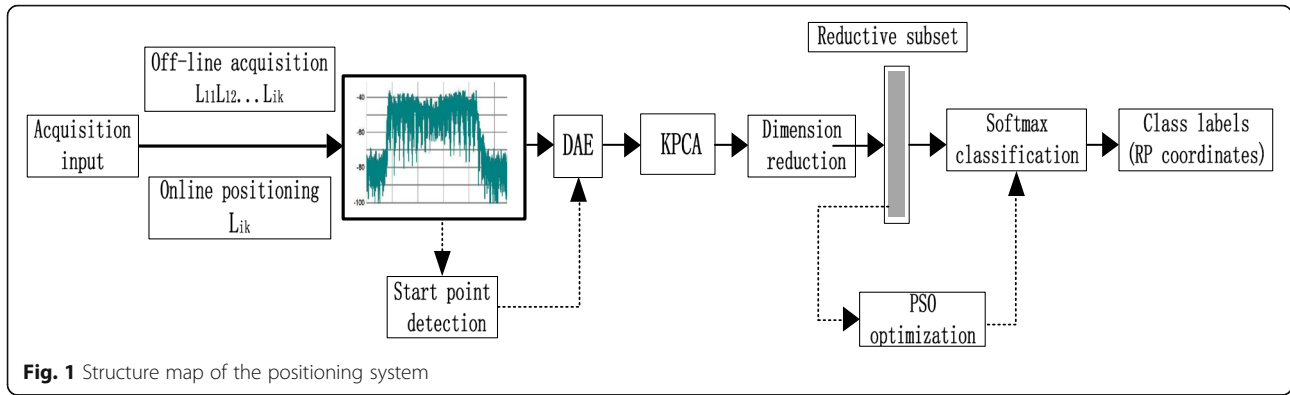


Fig. 1 Structure map of the positioning system

and steady-state spectrum characteristics. The former has a short duration and is difficult to capture. Therefore, steady-state spectrum acquisition is the main method in the research.

The experimental acquisition is carried out in the 20 m × 13 m open classroom of AI College. The RP coordinates are divided into equal intervals according to the grid. The SA44B spectrum receiver of Signal Corporation of America is deployed on the support of 0.8 m from the center, and the data collected by the receiver is transmitted to the server through the USB data line. The unlicensed illegal mobile phone communication jamming source is deployed on the same level of support of the receiver. Its essence is a modulation signal transmitter with controllable frequency band and mode. It can set up interference modulation mode of CDMA, GSM, DCS, and PHS. Its radio frequency range is 800–1990 MHz, 1 W transmitting power, and 2 dB transmitting antenna gain and can cover radius 30–40 m. The distance between the interference source and the receiver is determined by adjusting the bracket.

The spectrum receiver is set as SignalTrace mode at the system end, which can track the center frequency of the current jammer signal and output the spectrum to the interface. The sampling span SPAN is set to 140 Mhz, and the resolution bandwidth RBW is set to 100 kHz. In order to clearly compare the illegal signal with the normal signal, the jammer is set as DCS mode at the distance of 2 m between AP and RP. Start the detection program to collect the data of the receiver. Sampling time is 1.250 s, and 8561 data points are collected. The interference signals collected separately and the spectrum of the same frequency mobile phone dialing signals are superimposed and output to the interface, as shown in Fig. 2.

It can be seen from Fig. 2 that the spectrum of the radio signal can be seen:

The mobile phone signal is covered by the high-intensity signal of the interference source, and the spectrum bandwidth of the interference source is larger than the bandwidth of the mobile phone signal. The peak value of this type of interference source is not unique. Traditional indoor

fingerprint positioning methods mostly collect only one centroid frequency and corresponding RSSI values to represent the location characteristics of interference sources in Fig. 2. The spectrogram collected by the spectrum analyzer is rough and contains obvious noise, so it is difficult to extract the characteristics of the radio wave signal.

#### 2.4 Interference source signal pre-processing

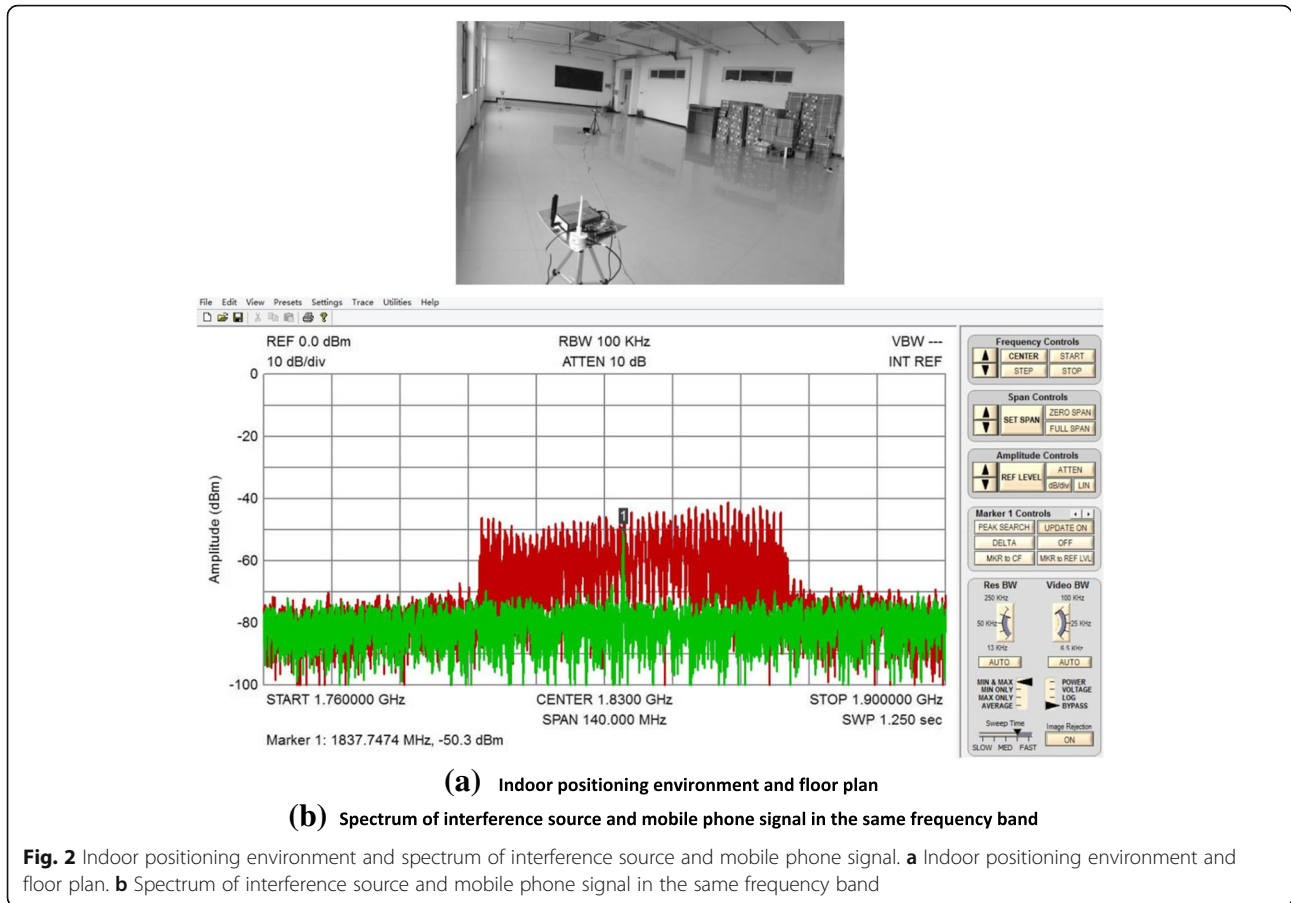
The rough edge of the signal spectrum will lead to too large error in extracting eigenvalues. By smoothing the noisy signal with the median filter, the sharp edge information can be well preserved. The smoothed spectrum with the median filter still contains noise, which needs to be denoised. In practical applications, the wavelet denoising method has a better effect on the spectrum Gauss white noise suppression. The commonly used wavelet denoising methods include modulus extremum denoising method [19], wavelet correlation denoising method [20], and wavelet threshold denoising method [21]. The commonly used wavelet threshold algorithms include the hard threshold algorithm and the soft threshold algorithm. However, the hard threshold wavelet denoising algorithm is prone to Gibbs oscillation. And the soft threshold wavelet denoising algorithm is prone to edge distortion due to the constant deviation of wavelet coefficients. A two-stage wavelet threshold function is constructed within the framework of Malat algorithm [22] to balance this problem. The traditional threshold function is improved as follows:

$$\widehat{W}_{\psi,\zeta} = \begin{cases} aW_{\psi,\zeta} + (1-a) \operatorname{sgn}(W_{\psi,\zeta}) \cdot (|W_{\psi,\zeta}| - b\lambda), & |W_{\psi,\zeta}| \geq \lambda \\ 0, & |W_{\psi,\zeta}| < \lambda \end{cases} \quad (1)$$

Among them, the primary threshold is set, and the secondary threshold is set according to the soft threshold function.

$$V = \frac{1}{0.6475} \sum |\delta_\rho^\psi| \sqrt{\delta_0};$$

The corresponding adjustment parameters are as follows:



$$a = \frac{\lambda}{W' + e^{|\lambda - \delta_\rho|}}, \quad b = \frac{|V - \delta_\rho|}{W' + e^{|\lambda - \delta_\rho|}};$$

In which  $W'$  is the mean of neighborhood wavelet coefficients, and the primary threshold corresponds to the “coarse tuning” signal. The adjustment parameter  $a$  decreases with the decrease of the wavelet coefficients in the range of  $(0, 1)$ . When the wavelet coefficients approach the threshold  $\lambda$ ,  $a \rightarrow 1$ , when the wavelet coefficients approach zero,  $a \rightarrow 0$ , and the wavelet coefficients are updated to  $\delta_1 = a_\rho^* \rho$  after the primary denoising. The secondary threshold  $b$  is independent according to the sparsity of sample points. The adjustment parameter corresponds to the “fine tune” signal, and the wavelet coefficient is updated to  $\delta_2 = \delta_\rho - b^* \delta_\rho$  after secondary noise reduction. After the two-level threshold thickness is adjusted to denoise the spectrum, the coefficients of each scale operation are used to reconstruct the signal. The API program of the spectrum analyzer is called by LabVIEW software, and combined with MATLAB programming, the median filter wavelet denoising of interference source input signal is processed, and the

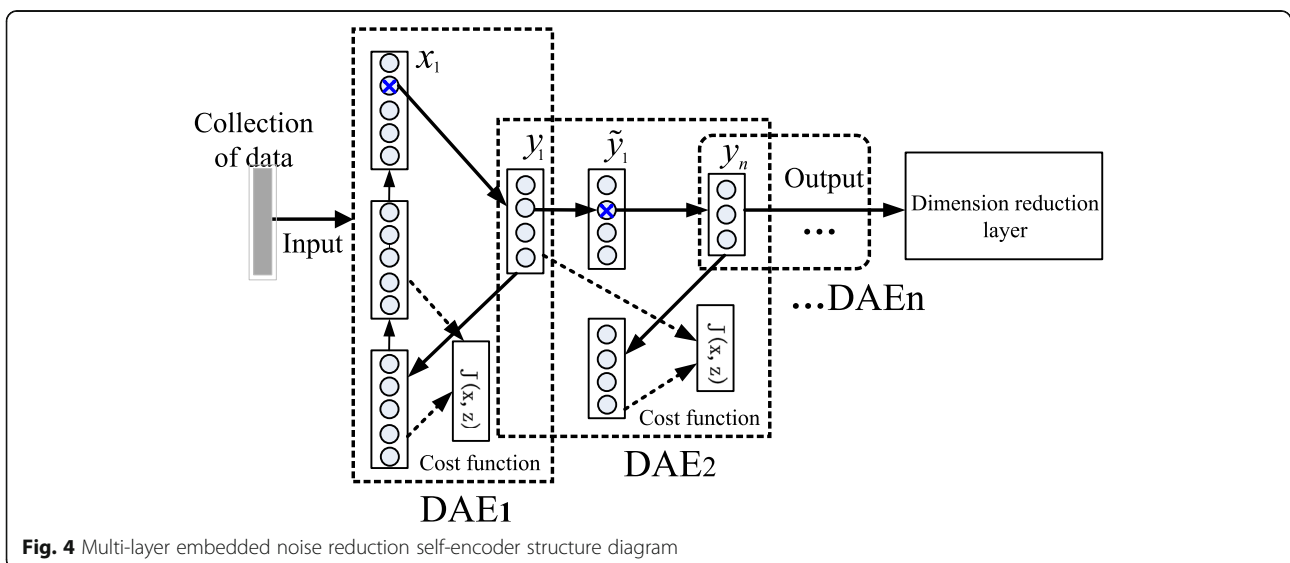
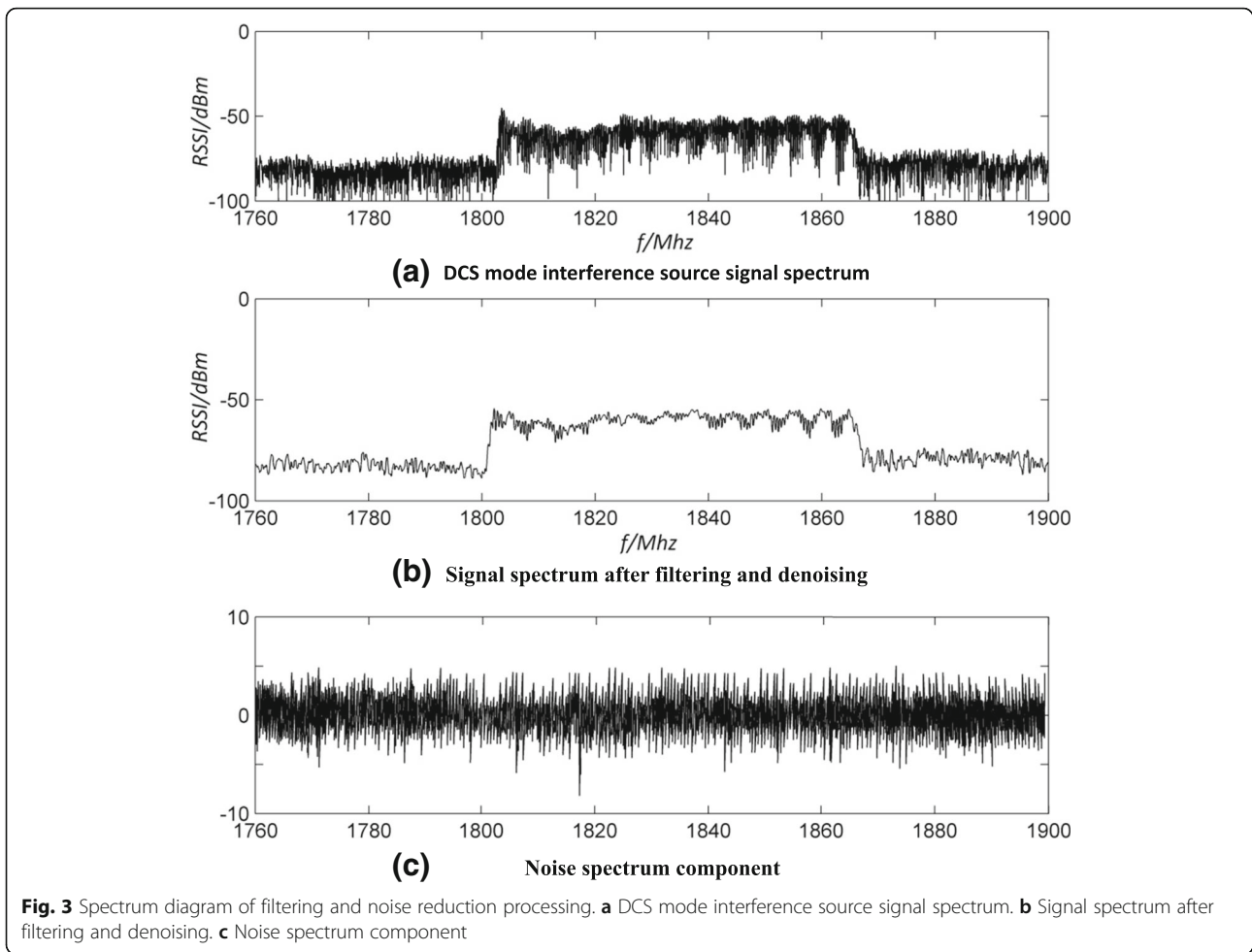
spectrum diagram of noise separation as shown in Fig. 3 is obtained.

### 3 Selection and extraction algorithms of spectrum signal eigenvalues

#### 3.1 Selection of signal characteristics

Traditional eigenvalue extraction mostly adopts manual extraction method, which is inefficient and unsuitable for optimal feature selection. In recent years, deep learning technology has been favored. Among them, DAE algorithm [23] belongs to unsupervised learning algorithm. It is based on the application of auto-encoder (AE) [24], which actively distributes unlabeled data into noise, and then trains. The training system acquires noise-free input by learning to remove noise, which is more robust to the training of random input signals. The core of the training system is the extension of the neural network algorithm.

The traditional multi-layer DAE algorithm is modified by artificially setting the number of layers to the adaptive parametric multi-layer embedded APADAE algorithm. As shown in Fig. 4, the front DAE1 output layer is linked to the front-end layer of the second DAE2 neuron, and the second DAE2 output is linked to the front-end of



the back-end neuron. The unlabeled sample data set  $x$  is mapped from the DAE1 layer to the hidden layer  $y$ , and the full connection layer is reconstituted by  $y$  to  $z = g(y) = s(w'y + b')$ . The minimum value of loss function  $L(x, y)$  constructed by  $x, y$ , which is obtained through parameter optimization. Among them,  $y = s(wx + b)$ ,  $w$  is the weight matrix of 1 dimension,  $w'$  is the constraint weight of self-encoder, and it is a reciprocal matrix with  $w$ , that is  $w' = w^T$ ,  $b$  is deviation function, and  $s$  is activation function. The  $s$  function is selected as the standard sigmoid function.

In this study, the data in the initial input layer of DAE1 vector  $x$  are “destroyed” by adding artificial noise. The data in  $x$  are randomly zeroed proportionally to form  $\tilde{x}_1$  as the input of the next layer of DAE2 and then transmitted to the reconstructed output layer  $\tilde{z}_2$  via the hidden layer  $\tilde{y}_2$  of DAE2 mapping. The optimal solution of the system can make the reconstructed error the most by adjusting the parameters of each layer. The cost function is defined as follows:

$$\begin{aligned} J(\omega, b) &= \arg \min \frac{1}{n} \sum_{i=1}^n L(\tilde{x}_n, \tilde{z}_n) \\ &= \arg \min \frac{1}{n} \sum_{i=1}^n L[\tilde{x}_n, g(f(\tilde{x}_n))] \end{aligned} \quad (2)$$

$L(\tilde{x}_n, \tilde{z}_n) = \|\tilde{x}_n - \tilde{z}_n\|^2$  is the loss function of APADAE network and further assumes that the input vector  $x$  obeys Bernoulli distribution, so the reconstructed cross-soil loss function is constructed as follows:

$$L(\tilde{x}_k, \tilde{z}_k) = - \sum_{k=1}^d [\tilde{x}_k \lg \tilde{z}_k + a^k (1 - \tilde{x}_k) \lg (1 - \tilde{z}_k)] \quad (3)$$

Among them, the latter adjusts the weight attenuation and controls the weight of cost function by adjusting the parameter  $a^{(k-1)}$  to avoid overfitting. Deep learning network (DLN) improves the output precision by feedback factor in its back propagation error algorithm. Residual  $\xi^{(k)}$  is defined to represent the difference between reconstructed  $\tilde{y}_k$  of DAE<sub>k</sub> layer network and high-dimensional input  $\tilde{z}_{k-1}$ . The layer  $k$  is determined by calculating the minimum loss function iteratively. First, we calculate the residuals  $\xi^{(k)}$  of the output DAE<sub>k</sub> layer:

$$\xi^{(k)} = - \left( y^{(k-1)} - x^{(k-1)} \right) \tilde{x}_k \left( y^{(k)} \right) \quad (4)$$

$$\begin{aligned} \tilde{x}_k \left( y^{(k)} \right) &= \tilde{x}_{k-1} \left( \xi^{(k)} \right) \left( 1 - \tilde{x}_{k-1} \left( \xi^{(k)} \right) \right) \\ &= \left( e^{y^{(k)}} + e^{-y^{(k)}} + 2 \right)^{-1} \end{aligned} \quad (5)$$

Then, iteratively calculate the residual  $\xi^{(k-1)}$  of the hidden layer. From the backpropagation error algorithm,

we can see that the descent vector along the gradient is as follows:

$$\frac{\partial J(\omega, b)}{\partial \omega} = \xi^{(k-1)} \left( a^{(k-1)} \right)^T \quad (6)$$

$$\frac{\partial J(\omega, b)}{\partial b} = \xi^{(k-1)} \quad (7)$$

When the above formula converges to the optimum value, the parameter vector  $\omega, b, a^{(k-1)}$  is used as the optimum adapting parameter of APADAE network model layer and is used to train the sample data. Sample data are input into multi-layer embedded DAE network one by one according to the algorithm flow to select features and map them to low-dimensional vector output.

Simplified experiments were conducted to verify the feature selection effect of APADAE algorithm. Six network output nodes were set up, i.e., six feature points selected from the single spectrum. The interference sources were deployed on eight RP coordinates one by one. One AP was fixed at the center of the classroom, and one spectrum was collected by each RP. Each RP label pair was extracted by APADAE algorithm selection. According to the characteristic data of the spectrum map, eight groups of features are extracted. As shown in Table 1, each row corresponds to a class label, each class label corresponds to six feature points, each feature point corresponds to two dimensions (signal strength  $RSSI$  and frequency  $f$ ), and each feature point corresponds to 12 dimensions. In the experiment, the parameter vector  $k$  of the algorithm APADAE maps to 11 layers of network.

### 3.2 Analysis of characteristic data

As shown in Table 1, it is obvious that unsupervised machine learning method has no prior autonomous feature extraction throughout the spectrum, and its dispersion is high and irregular. As the number of RP labels and AP sampling spectrum increases, its spatial dimension and total feature number will be geometrically doubled.

When the least coordinate labels are collected as  $s \cdot t = 8$  in the indoor interference source location experiment, then  $m = 4$  collection points are deployed. Each collection point collects  $n = 1$  signal spectrums, and each spectrum selects  $j = 4$  feature points. Each feature point corresponds to two parameters  $RSSI_1, f_1, RSSI_2, f_2, \dots, RSSI_j, f_j$ , the number is  $2j = 8$ . The feature vector set of each coordinate label can be constructed, and the total dimension number of feature vectors can be calculated as  $m \times n \times 2j = 4 \times 1 \times 8 = 32$ , and the total number of eigenvalues can reach by  $m \times n \times 2j \times 8 = 32 \times 8 = 256$ . With the increase of the number of coordinate labels, sampling points, and spectrum, the number of data dimensions and eigenvalues will increase sharply. When the parameters got  $s \cdot t = 24, x = 4, y = 3, j = 12$ , calculation

**Table 1** DAE algorithm feature selection simplifies experimental data

Class label	RSSI/(dBm)						f/(Mhz)					
	RSSI1	RSSI2	RSSI3	RSSI4	RSSI5	RSSI6	f1	f2	f3	f4	f5	f6
L <sub>11</sub>	-52.5	-50.2	-51.8	-49.1	-51.6	-53.9	1.803	1.767	1.851	1.785	1.845	1.829
L <sub>12</sub>	-50.2	-49.6	-50.2	-51.6	-50.3	-49.6	1.813	1.827	1.869	1.807	1.856	1.863
L <sub>13</sub>	-49.6	-48.3	-49.7	-48.5	-50.1	-50.5	1.821	1.837	1.758	1.812	1.846	1.829
L <sub>14</sub>	-50.5	-49.4	-51.3	-50.7	-51.5	-51.7	1.809	1.826	1.873	1.774	1.871	1.817
L <sub>21</sub>	-51.4	-51.5	-50.7	-49.3	-49.4	-50.5	1.858	1.842	1.845	1.856	1.855	1.829
L <sub>22</sub>	-43.9	-44.7	-43.5	-42.2	-44.5	-43.2	1.783	1.851	1.841	1.821	1.869	1.855
L <sub>23</sub>	-43.2	-44.1	-43.8	-44.7	-41.6	-42.5	1.861	1.844	1.856	1.809	1.841	1.828
L <sub>24</sub>	-50.6	-49.5	-49.8	-51.2	-49.6	-50.3	1.834	1.861	1.863	1.845	1.857	1.864

shows that the overall feature dimension contains  $m \times n \times 2j = 4 \times 3 \times 24 = 288$ , and the total quantity is up to  $m \times n \times 2j \times 24 = 288 \times 24 = 6912$  Fig. 5.

### 3.3 Dimension reduction

From the analysis of feature data, it can be seen that the increase of data will lead to too much computational complexity, which may lead to “dimension disaster,” and slow data training process will affect generalization ability. Moreover, the feature vectors selected by APADAE algorithm are not all effective, so the redundant feature vectors are eliminated without reducing the classification accuracy, and a feature reduction subset is constructed. The main methods of dimension reduction are compressed sensing, principal component analysis, convolution mapping, and threshold filtering.

KPCA is a commonly used nonlinear dimension reduction method [25]. It is based on the kernel function technique to “kernelize” the linear dimension reduction method. It can be used to improve the ability of principal component analysis (PCA) to analyze nonlinear sample data. Mercer kernel function is introduced into the study, and evaluation parameters are embedded to adjust the proposed method. In order to reduce the dimension of feature data and eliminate the multiple correlations between vectors, the component process is used to extract the high-dimensional nonlinear reduction subset with high cumulative contribution rate.

The feature space data  $Z = (z_1, z_2, \dots, z_j)$  extracted from the APADAE algorithm is standardized by  $\tilde{z}_i = (z_i - \mu_i) / \sigma_i$ ,  $\mu$  is the sample mean, and  $\sigma$  is the sample standard deviation, and then, it is mapped to the hyperplane of  $W = (w_1, w_2, \dots, w_d)$ .

$$\left( \sum_{i=1}^m v_i v_i^T \right) w_j = \ell_j w_j \quad (8)$$

$v_i$  is the mapping of a sample point  $z_i$  in high-dimensional feature space, so we got the parameter  $w_j$ .

$$w_j = \frac{1}{\ell_j} \left( \sum_{i=1}^m v_i v_i^T \right) w_j = \sum_{i=1}^m v_i \frac{v_i^T w_j}{\ell_j} = \sum_{i=1}^m v_i \alpha_i^j \quad (9)$$

$\alpha_i^j = \frac{1}{\ell_j} v_i^T w_j$  is the  $j$  component of  $\alpha_i$ , and  $w$  is transformed into  $w_j = \sum_{i=1}^m \phi(\tau_i) \alpha_i^j$ . The Mercer kernel function is introduced as follows:

$$\kappa(\tau_i, \tau_j) = \phi(\tau_i)^T \phi(\tau_j) \quad (10)$$

That is  $K\alpha^j = \ell_j \alpha^j$ , and  $K$  is the kernel of  $\kappa$  corresponding to the nuclear parameter  $\tau$ . It can be got that  $K(\tau_i, \tau_j) = \kappa(\tau_i, \tau_j)$  and  $\alpha^j = (\alpha_1^j; \alpha_2^j; \dots; \alpha_m^j)$ . According to eigenvalue decomposition, the largest eigenvalue  $\ell$  of  $K$  and its corresponding eigenvectors are obtained.

$$v_j = w_j^T \phi(\tau) = \sum_{i=1}^m \alpha_i^j \kappa(\tau_i, \tau) \quad (11)$$

In (11), the  $j$  dimension coordinate vector after the mapping of the new data point  $\tau$ .

The selection of kernels has great influence on the implementation of KPCA. The common kernels are linear kernels, polynomial kernels, Gauss radial basis function, sigmoid kernels, etc. Gauss Radial Basis Kernel (RBF) function is easy to extract local features from different kinds of labels while retaining the overall feature information. Therefore, it is chosen as a nonlinear relationship between the eigenvalues to mine. Its mathematical expression is as follows:

$$K(\tau_i, \tau_j) = \exp\left(\frac{|\tau_i - \tau_j|^2}{-\varepsilon^2}\right) \quad (12)$$

In (12), the  $\varepsilon$  is the radial base nucleus width. In order to measure the interclass separability, we define the hyper parameters  $d_L$  to adjust the extraction process of the  $L$  principal component (class label) process.

$$d_L = \frac{1}{L(L+1)} \sum_{\alpha=1}^{L-1} \sum_{\beta=\alpha+1}^L \left( 1 - \exp \left[ -\frac{d_{\alpha\beta}}{r_{\alpha}^L + r_{\beta}^L} \right] \right) \quad (13)$$

In (13), we can get  $d_{\alpha\beta}$  is the  $L$  principal classification intervals between the  $\alpha$  and the  $\beta$  class label. Then,  $r_{\alpha}^L$  and  $r_{\beta}^L$  are the maximum distance between the  $E$  principal component and the center of class  $\alpha$  and  $\beta$ , respectively. And  $\eta = |(d_{\alpha} - d_{\beta})|/d_L$  is defined as the adjustment margin of  $d_L$ . When  $\eta$  reaches the minimum value,  $d_L$  takes the maximum value.

The ratio of extracted principal components to the cumulative variance contribution rate of eigenvalues is weighed.

$$P_{CPV} = \sum_{j=1}^i \ell_j / \sum_{j=1}^n \ell_j \quad (14)$$

The value of  $P_{CPV}$  represents the size of principal component implication information. In order to prevent data from overfitting, we select feature vectors of  $P_{CPV}$  value before  $c$  and construct special reduction sets  $c = 2$ .

The implementation steps of KPCA dimension reduction algorithm are the following:

Step 1: Initialize the program, normalize the sample data to get the mapping matrix  $W$ , and then get the Gauss kernel matrix  $K$ ;

Step 2: The eigenvalues and eigenvectors are calculated, and the eigenvalues of the correlation coefficient matrix are sorted  $\ell_1 \gg \ell_2 \gg \dots \gg \ell_j$ , and the eigenvector  $v_1, v_2, \dots, v_j$  corresponding to Schmidt orthogonalization is obtained;

Step 3: The  $d_L$  and  $P_{CPV}$  are evaluated and evaluated.

Step 4: When  $\eta$  approaches 0.0001, the program terminates, and the feature vectors corresponding to the principal components of  $P_{CPV}$  before  $c$  sort are output to form a new reduced feature set.

### 3.4 Feature vector extraction

After processing Table 1 data by the KPCA algorithm, where the value of the core width parameter  $\varepsilon$  is equal to 0.5, we get the eigenvalue  $\ell$  and the cumulative contribution rate  $P_{CPV}$  of the data, and the order is as follows:

Comparing with Table 2, it is obvious that the KPCA algorithm extracts eigenvector components more significantly than PCA. The PCA algorithm extracts eight vectors according to approximate values, while the KPCA algorithm can extract 14 eigenvectors with stronger feature representation ability. The first four of them contain 95.81% of the original signal feature information, but the computational overhead of KPCA algorithm is slightly larger. The feature vectors of the first four ranking values of KPCA algorithm are selected as the feature

**Table 2** The data of two PCA methods

Principal component ordination	PCA		KPCA	
	$\ell$	$P_{CPV}/\%$	$\ell$	$P_{CPV}/\%$
1	0.628 2	51.53	0.394 8	52.72
2	0.314 7	25.82	0.116 6	25.72
3	0.137 5	11.28	0.069 3	10.59
4	0.075 8	6.22	0.050 8	6.78
5	0.039 1	3.21	0.046 7	5.17
6	0.016 1	1.32	0.029 5	2.74
7	0.007 4	0.61	0.015 3	1.78
8	0.000 2	0.02	0.009 8	1.31
9	0	0	0.006 2	0.83
10	0	0	0.004 3	0.57
11	0	0	0.003 5	0.47
12	0	0	0.001 6	0.21
13	0	0	0.000 4	0.05
14	0	0	0.000 1	0.01
15	0	0	0	0
16	0	0	0	0

representation of the coordinate label, that is, the original 16-dimensional feature vectors are replaced by four dimension optimal feature vectors, and the dimension reduction is 1/4 and the labeled optimal feature set is formed.

## 4 Classifier construction, optimization, and training test

The reduced subset after KPCA processing is classified and processed. The number of classes is equal to the label number or the number of RP coordinates. There are many commonly used shallow learning classification methods: the BP neural network, the support vector machine (SVM) classification, the Softmax classification [26], and so on. Most of the binary classification methods need to change and combine to form a multi-classification pattern. When the dimension is too large, the increase of the classification level results in the slow process of the algorithm.

### 4.1 The establishment of classifier

According to the analysis of APADAE network algorithm, it is obvious that not only dimensionality reduction but also classification function can be achieved by controlling the number of nodes in the output layer, but the extraction category is rough and inefficient. Softmax classification model, which is widely used in the field of in-depth learning, and PSO optimization process are introduced to improve classification efficiency. The core of this research is to normalize the mapping values of



neurons in the classification layer into class label probabilities, which is essentially an extension of the logistic regression model in multi-classification applications.

Defining the feature subset  $U$  is  $\{u^{(1)}, u^{(2)}, \dots, u^{(l)}\}$ ,  $l \in [1, L]$ , the probability of input samples according to corresponding class labels is  $p(L^{(l)} = L | u^{(l)}; \theta)$  and to realize Softmax-based multi-class classification, its logarithmic regression loss function is defined as:

$$h_{\theta}(u^{(l)}) = \begin{bmatrix} p(L^{(l)} = 1 | u^{(l)}; \theta_1) \\ p(L^{(l)} = 2 | u^{(l)}; \theta_2) \\ \vdots \\ p(L^{(l)} = L | u^{(l)}; \theta_L) \end{bmatrix} = \frac{1}{\sum_{l=1}^L e^{\theta_l^T \cdot u^{(l)}}} \begin{bmatrix} e^{\theta_1^T \cdot u^{(l)}} \\ e^{\theta_2^T \cdot u^{(l)}} \\ \vdots \\ e^{\theta_L^T \cdot u^{(l)}} \end{bmatrix} \quad (15)$$

In the above formula,  $L$  is the number of class labels, and system parameter  $\theta$  determines the probability of class labels and here assigns  $\theta = (\theta_1^T, \theta_2^T, \dots, \theta_L^T)$ . Matrix  $\theta$  corresponds to the probability of one class label per row. The maximum probability  $p(L^{(l)} = L | u^{(l)}; \theta)$  is the classification label of class  $u^{(l)}$  corresponding to  $l$  class, and the update cost function is

$$Q(\theta) = -\frac{1}{2j} \left[ \sum_{l=1}^{2j} \sum_{l=0}^L 1\{u^{(l)} = l\} \log \frac{e^{\theta_l^T u^{(l)}}}{\sum_{l=1}^L e^{\theta_l^T u^{(l)}}} \right] \quad (16)$$

The gradient derivative method is applied to solve the partial derivatives of the classification labels, and the corresponding class label parameter  $\theta_L$  is obtained by the following formula.

$$\frac{\partial Q(\theta)}{\partial \theta_L} = -\frac{1}{2j} \left[ \sum_{l=1}^{2j} u^{(l)} \left( 1\{L^{(l)} = l\} - P(L^{(l)} = l | u^{(l)}; \theta) \right) \right] \quad \begin{cases} P_r = TP / (TP + FP) \\ R_e = TP / (TP + FN) \\ F_s = 2P_r R_e / (P_r + R_e) \end{cases} \quad (18)$$

#### 4.2 PSO optimizes the classification process

As an evolutionary parallel algorithm, PSO algorithm [27] is widely used because its particles traverse the space iteratively from random solutions and have high precision and fast convergence. PSO is used to optimize the classification model of Softmax in order to speed up the classification process. The steps are as follows:

Step 1: Constructing the cost function of Softmax multi-classification model.

Step 2: The location of particles is initialized in the vector space, where the initial dimension of the particle is 16 and the number of iterations is 100.

Step 3: The fitness function  $Q(\theta)$  is adjusted by gradient descent method, and the parameter  $\theta_L$  of fitness is verified by  $K$ -fold cross-validation. The fold is set to 5. The fitness value  $Q(\theta_l)$  is iteratively calculated to update the position of particles to achieve global optimization.

Step 4: When the threshold condition is reached, the optimal parameter  $\theta_L$  of Softmax can be obtained by outputting the position parameters of the particles.

Step 5: In order to prevent the problem of "gradient cliff" in convex optimization processing, the Polyak averaging algorithm is applied to process the output of  $\theta_L$  by  $q = 100$  iterations to get  $\hat{\theta}_L = q^{-1} \sum_l \theta_L^q$ , and then, the label corresponding to the probability of sample classification is obtained.

#### 4.3 Training test validation

According to the above strategy of the off-line phase, 8, 12, 16, and 24 RP coordinates were selected respectively as the experimental environment organized in section 1.1, and a spectrum analyzer was collected successively around RP points on four different AP coordinates. Each RP coordinate in DCS mode of interference source sampled three groups of spectrum separately. In order to prevent the instantaneous overshoot phenomenon of acquisition signal from lasting for 20 s, the acquisition time interval was 10 s to extract the steady-state spectrum. The total number of 24, 36, 48, and 72 groups of spectrum were collected respectively. In order to prevent data overtraining, 1/6 of the collected data was randomly selected for classifier training and parameter optimization. The remaining data was used for classifier testing. The performance of classification testing was measured by precision ( $P_r$ ), recall rate ( $R_r$ ), and F1 score (F1-score,  $F_s$ ) [28].

In (18), the four parameters TP, TN, FP, and FN correspond to the true rate, the true negative rate, the false positive rate, and the false negative rate of the resolution results, respectively.

#### 5 Results and discussion

We collect data from the same environment in Section 2. The classroom is open and barrier-free. The interference source and the shelf are placed at the center coordinate (9.6, 6.4 m), and the slurry antenna is placed vertically. We collect the spectrum data of interference

sources one by one from the edge of the classroom by a single receiving AP according to the grid. The width of the mesh is 0.8 m and 1.2 m, respectively. The AP moving position coordinates are as follows: (2.5, 1.2), (4.5, 1.2), (6.5, 1.2) et al., and the RP position is (2, 1.2), (4, 1.2, 6, 1.2) and so on. In addition, we changed the number of RPs at the same acquisition interval. This makes the dimension of data computation increase from 32 to 288. The experiment is to verify the correlation between dimension and positioning accuracy by changing dimension parameters. We delete the location coordinate data which are far away or unsuitable for acquisition and classify and evaluate them using the method of Softmax (Fig. 5). In addition, we compare it with the literature BP [29] and SVM [30] in Table 3.

From Table 3, it can be seen that when the number of RP is 8, and the grid of RP the location interval distance is selected as one eighth of the width of the test classroom, that the grid of RP is about 1.625 m. When the number parameter of RP increases to the maximum of 24, the interval distance is equal to 24 equal parts of the width of the test room, which is approximately 0.541 m. When RP equals 8, we can get three evaluation values of  $P_r$ ,  $R_r$ , and  $F_s$ , which correspond to three algorithms of BP, SVM, and Softmax. Comparatively speaking, BP algorithm obtains higher PR value because of its higher computational complexity than SVM algorithm. The highest PR value of Softmax algorithm is 80.2. It can be seen that the training scores of the three classification methods are better in Softmax, followed by BP, and SVM needs to cooperate with a multi-classification strategy to make its score the lowest. Obviously, the more to add the number of RP, the higher to get the performance. Considering the limitation of the experimental

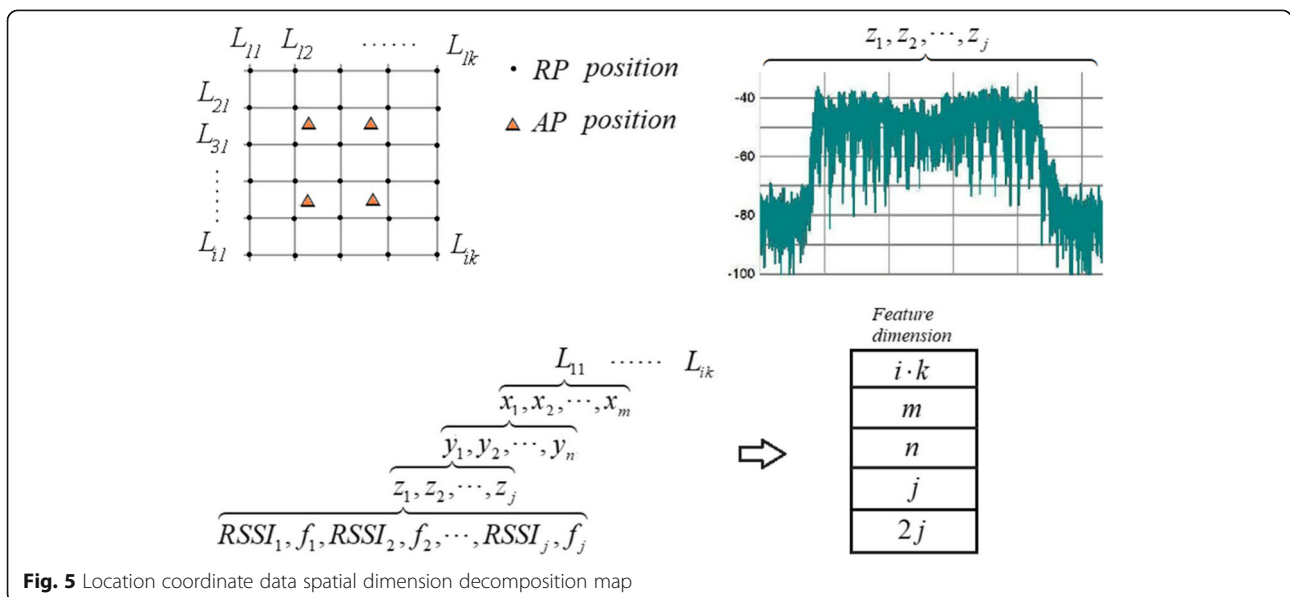
**Table 3** RP number and feature classification performance

RP number	BP			SVM			Softmax		
	$P_r$	$R_r$	$F_s$	$P_r$	$R_r$	$F_s$	$P_r$	$R_r$	$F_s$
8	73.8	66.2	69.7	63.1	58.8	60.8	80.2	63.6	70.9
12	76.7	65.7	70.7	65.8	61.3	63.4	83.4	65.7	73.5
16	78.3	60.5	68.2	72.6	65.7	68.9	91.1	89.8	90.4
24	77.6	64.3	70.3	76.1	63.8	69.4	86.8	71.9	78.7

environment and the operation cost, the RP selection 16 is relatively optimal.

In the further experiment, RP is equal to 16 and the first three of all classifiers are selected. The interference sources are randomly placed in 32 non-label locations in the room to observe the relationship between resolution accuracy and location error. Five groups of spectrum are detected at each location. The classification results of Softmax and the classification errors of three algorithms are listed in Fig. 6 and Table 4 in terms of space relationship.

It is very obvious that the data in Fig. 6 are regular. It is a comparison result of position labels, in which the horizontal and vertical coordinates are from  $L_{11}$  to  $L_{44}$ . This is the data collected when the RP parameter is selected as 16, which the grid of RP is about 0.813 m. The values of each row and column correspond to the actual position coordinates of the test classroom environment. We can clearly see that the diagonal data is the maximum probability of locating. For example, we get 91.6 values at  $L_{11}$ , and 5.8 and 2.6 values at  $L_{12}$  and  $L_{14}$ , respectively. This means that when the reference point is selected at  $L_{11}$ , about 5 of the 100 measurements are positioned at  $L_{12}$  and 2 at  $L_{14}$ . Obviously, this is the positioning error. We use the algorithm to express the



**Fig. 5** Location coordinate data spatial dimension decomposition map

	L <sub>11</sub>	L <sub>12</sub>	L <sub>13</sub>	L <sub>14</sub>	L <sub>21</sub>	L <sub>22</sub>	L <sub>23</sub>	L <sub>24</sub>	L <sub>31</sub>	L <sub>32</sub>	L <sub>33</sub>	L <sub>34</sub>	L <sub>41</sub>	L <sub>42</sub>	L <sub>43</sub>	L <sub>44</sub>
L <sub>11</sub>	91.6	5.8		2.6												
L <sub>12</sub>	7.1	89.4				3.5										
L <sub>13</sub>		7.9	92.1													
L <sub>14</sub>				91.3	8.7											
L <sub>21</sub>	3.9				91.5	4.6										
L <sub>22</sub>		3.3				93.3				3.4						
L <sub>23</sub>						4.8	95.2									
L <sub>24</sub>							2.7	90.6	6.7							
L <sub>31</sub>					3.3				91.6	2.8			2.3			
L <sub>32</sub>						2.6			2.8	92.7	1.9					
L <sub>33</sub>										2.1	91.7	2.1			4.1	
L <sub>34</sub>								2.4			3.2	92.2				2.2
L <sub>41</sub>									3.3				91.4	5.3		
L <sub>42</sub>													7.1	89.6	3.3	
L <sub>43</sub>											1.8			4.8	88.3	5.1
L <sub>44</sub>												6.4			2.2	91.4

Fig. 6 Feature vector classification results

positioning error clearly by comparing the probability. In a word, the data of other coordinates are the same, and the complete positioning error data of the test environment can be obtained through Fig. 6.

As can be seen from the test results of mixed matrix and Table 4 in Fig. 6, there are 16 diagonal matrices for correct classification. The overall correct average recognition rate can reach 91.5%, and the overall average positioning accuracy can reach 1.16 m. Softmax classification is more effective than other classifiers in identifying the RP position of interference sources. The recognition rate and positioning accuracy are basically positively correlated. At the same time, it is also observed that in the middle area of the location environment, the identification rate is high, up to 95.2%. And the resolution errors mostly occur in the corner coordinates of the room, where the signal refraction causes great fluctuation, which confirms that indoor environment has a greater impact on location.

Table 4 Positioning error

	BP	SVM	Softmax
Average error/m	1.41	1.23	1.16
Average recognition rate/%	73.4	78.8	91.5

### 6 Conclusion

By changing the traditional fingerprint positioning method of interference source, a new positioning system is established, which verifies the positioning accuracy and reliability of the new indoor positioning method of interference source. The method of extracting spectrum characteristic parameters of indoor interference sources is effective. The average recognition rate of feature can reach 91.5%, and the average positioning accuracy is 1.16 m, which compared with the previous positioning methods is greatly improved. As to the positioning environment in the experiment, the number of samples collected is not positively related to the positioning accuracy, and the 192 dimension of the overall dimension is close to the optimal.

The experimental model in this study can provide a new reference for the realization of indoor jamming source location. When the types of jamming sources and ranging and jamming communication modes are changed, their signal characteristics will change accordingly. This requires a large number of real-time updates of sample data to enrich the location dictionary. In subsequent experiments and studies, it will be gradually improved.

### Abbreviations

APADAE: Adaptive parameter adjustment denoising auto-encoder algorithm; DAE: Deep learning, denoising auto encoder; KPCA: Kernelized principle component analysis; PSO: The particle swarm optimization algorithm; SVM: The support vector machine

### Acknowledgements

Not applicable.

### Funding

This paper is supported by the National Natural Science Foundation of China (F2014202264) and the Cooperation Fund of Ministry of industry and information technology (12-MC-KY-14).

### Availability of data and materials

Data sharing not applicable to this article as the data ownership attributable to Radio Monitoring Committee of Beijing.

### Authors' contributions

YC carried out the spectrum studies and participated in drafting the manuscript. TD conceived of the study, participated in its design and coordination, and reviewed the work. CJ carried out the design of the algorithms and participated in the indoor location experiments. SS participated in the design of the study and performed the data analysis. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 November 2018 Accepted: 29 January 2019

Published online: 26 February 2019

### References

1. X. Li, E. Björnson, E.G. Larsson, S. Zhou, W. Jing, Massive mimo with multi-cell mmse processing: exploiting all pilots for interference suppression. *EURASIP J. Wireless Commun. Netw.* **2017**(1), 117 (2017)
2. Z. DENG, Y. YU, X. YUAN, N. WAN, L. YANG, Situation and development tendency of indoor positioning. *China Commun.* **10**(3), 42–55 (2013)
3. G.H. Huff, J. Feng, S. Zhang, J.T. Bernhard, A novel radiation pattern and frequency reconfigurable single turn square spiral microstrip antenna. *IEEE Microw. Wireless Compon. Lett.* **13**(2), 57–59 (2015)
4. M.H. Zhang, Received-signal-strength-based indoor location in wireless LANs. *Comput. Sci.* **34**(6), 68–71 (2007)
5. H. Guo, N. Reisi, W. Jiang, W. Luo, Soft combination for cooperative spectrum sensing in fading channels. *IEEE Access* **5**, 975–986 (2017)
6. L. Mingxin, S. Jianli, Design and implementation of WLAN indoor positioning system model based on energy efficiency. *Chin. J. Sci. Instrum.* **5**(35), 1169–1178 (2014)
7. M. Yin, J. Gao, Z. Lin, Laplacian regularized low-rank representation and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(3), 504–517 (2016)
8. M. Wu, Y. Wu, L. Xiao, M. Ming, A. Liu, Z. Ming, Learning-based synchronous approach from forwarding nodes to reduce the delay for industrial internet of things. *EURASIP J. Wireless Commun. Netw.* **2018**(1), 10 (2018)
9. V. Sharma, M. Bennis, R. Kumar, Uav-assisted heterogeneous networks for capacity enhancement. *IEEE Commun. Lett.* **20**(6), 1207–1210 (2016)
10. R. Zhu, M. Ma, L. Liu, S. Mao, Cooperative and intelligent sensing: a special section in IEEE access. *IEEE Access* **5**, 27824–27826 (2017)
11. A. Boulch, P. Trouvé, E. Koeniguer, F. Janez, B. Le Saux, *Learning speckle suppression in SAR images without ground truth: application to sentinel-1 time-series*, Geoscience and Remote Sensing Symposium IGARSS 2018–2018 IEEE International (2018), pp. 2366–2369
12. A.Z. Papafragkakis, A.D. Panagopoulos, *Impulse radio ultra wideband for broadband indoor machine-to-machine communication networks: system performance evaluation[C]// international workshop on antenna technology: small antennas, innovative structures, and applications. IEEE* (2017)
13. J. Wang, R. Zhu, S. Liu, A differentially private unscented Kalman filter for streaming data in IoT. *IEEE Access* **6**, 6487–6495 (2018)
14. C. Bin, Z. Xi'an, A hybrid indoor positioning method based on propagation model and location fingerprint. *Bull. Surveying Mapp.* **6**, 35–38 (2015)
15. Y. Jia, J. Ma, L. Gan, Combined optimization of feature reduction and classification for radiometric identification. *IEEE Signal Process. Lett.* **24**(5), 584–588 (2017)
16. C. Xu, K. Li, Cooperative test scheduling of 3d noc under multiple constraints based on the particle swarm optimization algorithm. *Chin. J. Sci. Instrum.* **38**(3), 765–772 (2017)
17. Y. He, W. Meng, L. Ma, et al., *Rapid deployment of APs in WLAN indoor positioning system[C]//International ICST Conference on Communications and NETWORKING in China. IEEE Comput. Soc* (2011), pp. 268–273
18. I. Dey, P.S. Rossi, Probability of outage due to self-interference in indoor wireless environments. *IEEE Commun. Lett.* **21**(1), 8–11 (2017)
19. H. Liu, C. Zhao, L. Xuan, et al., Study on a neural network optimization algorithm based on improved genetic algorithm. *Chin. J. Sci. Instrum.* **7**, 1573–1580 (2016)
20. V.E. Kosmidou, P.C. Petrantonakis, L.J. Hadjileontiadis, Enhanced sign language recognition using weighted intrinsic-mode entropy and signer's level of deafness. *IEEE Trans. Syst. Man Cybern. B Cybern.* **41**(6), 1531 (2011)
21. X. Liu, R. Zhu, B. Jalaian, Y. Sun, Dynamic spectrum access algorithm based on game theory in cognitive radio networks. *ACM MONET* **20**(6), 817–827 (2015)
22. Z. Huang, G. Shan, J. Cheng, J. Sun, TRec: an efficient recommendation system for hunting passengers with deep neural networks. *Neural Comput. & Applic.* (2018). <https://doi.org/10.1007/s00521-018-3728-2>
23. M.J. Ho, S.M. Berber, K.W. Sowerby, Indoor cognitive radio operation within the broadcast tv protection contour. *Phys. Commun.* **23**, 43–55 (2017)
24. W. Liu, G. Liu, X. Ji, J. Zhai, Y. Dai, Sound texture generative model guided by a lossless Mel-frequency convolutional neural network. *Access IEEE* **6**, 48030–48041 (2018)
25. F. Gontier, M. Lagrange, P. Aumond, A. Can, C. Lavandier, An efficient audio coding scheme for quantitative and qualitative large scale acoustic monitoring using the sensor grid approach. *Sensors* **17**, 2758 (2017)
26. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Recent advances in convolutional neural networks. *arXiv:1512.07108* (2015)
27. H. Nam, M. Choi, S. Han, C. Kim, S. Choi, D. Hong, A new filter-bank multicarrier system with two prototype filters for QAM symbols transmission and reception. *IEEE Trans. Wirel. Commun.* **15**(9), 5998–6009 (2016)
28. C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, L. Hanzo, Machine learning paradigms for next-generation wireless networks. *IEEE Wirel. Commun.* **24**(2), 98–105 (Apr. 2017)
29. O. Steven Eyobu, D. Han, Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network. *Sensors* **18**, 2892 (2018)
30. A.I. Préz-Neira et al., MIMO signal processing in offset-QAM based filter bank multicarrier systems. *IEEE Trans. Signal Process.* **64**(21), 5733–5762 (2016)

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)