Applied Intelligence (2023) 53:9089–9102
https://doi.org/10.1007/s10489-022-03925-2

A generalized thresholding algorithm with dimension reduction for device-free localization in IoT

Qin Cheng[1,2] · Linghua Zhang[1] · Bo Xue[2,3] · Feng Shu[1] · Xu Wang[1]

Accepted: 22 June 2022 / Published online: 5 August 2022

Abstract
Location information is one of the most important factors for many location-based services (LBSs) in the Internet of Things (IoT). Device-free localization (DFL) has received more attention as it achieves localization without attaching any electronic device to the target. DFL can be applied to many special scenarios, such as monitoring the elderly living alone, health care of inpatients, and emergency rescue. In applications based on traditional localization methods, the numerous receive signal strength (RSS) measurements are collected from wireless sensor networks (WSNs) comprised of sensor pairs to construct the atoms of learning dictionaries. With recovery algorithms, solutions can be obtained from undetermined equations using learning dictionaries, which can be mapped to the position index of the target to estimate the accurate coordinates. However, the numerous RSS data produced by WSN sensor generate high-dimensional learning dictionaries that cost the sparse recovery algorithm more iterative computation time to derive the target location and more space for data storage, thus affecting the real-time DFL performance. In this paper, we propose a data dimension reduction method based on the generalized iterative thresholding algorithm for DFL. Firstly, we reduced the column and row dimensions of the dictionary, respectively, via principal components analysis (PCA). Then, the dimension of the observed vector was reduced correspondingly. Finally, the new underdetermined equation was solved via sparse coding with an iterative p-thresholding algorithm in signal subspace, and the target location was estimated accurately. Experiments on public datasets demonstrated that the proposed method outperforms the current alternatives by improving the computation efficiency of DFL systems and taking less time to locate the target, implying its good applicability to IoT scenarios with high real-time requirements.

Keywords Device-free localization · p generalized thresholding function · Dimension reduction · Principal component analysis

1 Introduction

Target location is one of the most fundamental pieces of information in the Internet of Things (IoT). Traditional positioning technologies include the global positioning system (GPS), Bluetooth, ultrasonic, and radio frequency identification (RFID) [1], which often require electronic equipment installed on the targets, such as electronic identification tags and wireless sensor nodes. Hence the name device-based localization (DBL). However, DBL cannot be applied to targets unwilling or unable to carry those devices, e.g., the elderly or invaders. Device-free localization (DFL) emerges as promising technology since it does not require any device and can be widely adopted in many special applications with space constraints, such as elderly behavior recognition in smart homes, target detection in disaster rescue, and indoor invader surveillance in security defends, as shown in Fig. 1.

To achieve DFL, multiple radio frequency (RF) sensors have been installed in advance to detect, locate and track targets in the area of interest. A typical DFL system consists of numerous pre-installed wireless sensor nodes, i.e., access points (APs), which act as transceivers to transmit and receive radio signals. When the target equivalent to an obstacle is at a different position in the area, the radio signal links passing through it are attenuated. Thus, the target location can be determined by detecting and estimating the weak signal changes.

Youssef et al. [2] first conceptualized device-free passive localization and formulated DFL as a fingerprint matching

✉ Linghua Zhang
zhanglh@njupt.edu.cn

Extended author information available on the last page of the article.

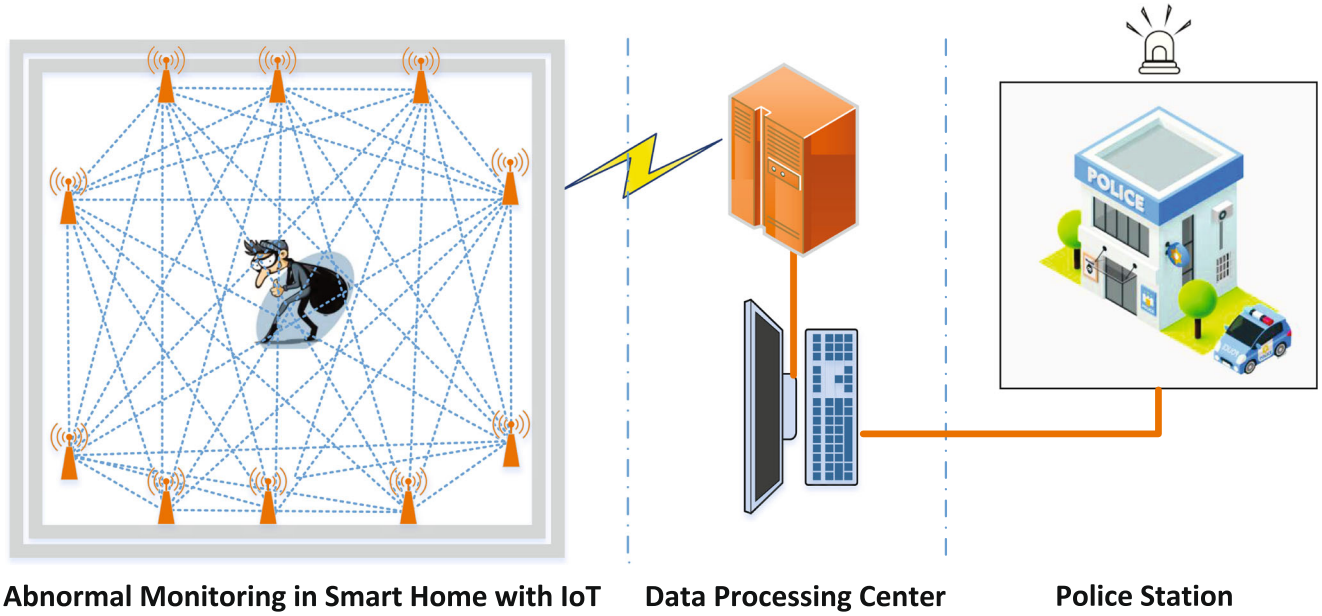

Fig. 1 Illustration of a DFL system applied to security monitoring in a smart home with IoT

problem. Zhang et al. [3] proposed three geometric tracking approaches for eliminating noise and improving the tracking accuracy of DFL systems, i.e., midpoint, intersection, and best-cover algorithms. To solve the ill-posed inverse problem, Wilson and Patwari [4] regarded DFL as an inverse problem, for which they presented a linear model and utilized Tikhonov regularization to obtain resultant images for DFL by radio tomographic imaging (RTI). The above studies provided an early foundation for DFL research.

Related works inspired by image processing methods in computer vision and pattern recognition imply the feasibility of classification methods in solving localization problems, e.g., deep neural networks with autoencoders [5, 6]. Among the different classification methods, sparse coding has attracted the attention of many researchers due to its simple decision rules, high accuracy and high efficiency [7]. Localization can be transformed into sparse representation classification (SRC) problems for DFL systems [8]. In other words, sparse solutions of underdetermined systems can be reached by sparse coding algorithms with sparsity constraints, and the target location can be recovered by seeking the relationship between the solution and the index of the preset monitoring area grids occupied by the target [9–12].

Solving the DFL problem via sparse coding is a process of inferring the target location from the sparse coefficient solutions obtained from underdetermined systems with observed signals and base dictionaries. Recently explored approaches to DFL via sparse coding included orthogonal matching pursuit (OMP) [10, 13, 14], basis pursuit linear programming (BP-LP) [8, 14], and iterative shrinkage-thresholding algorithm (ISTA) [11, 15–17]. However, BP-LP [8] lacks efficiency of computations, and the OMP algorithm in [10] cannot provide sufficient DFL accuracy. Li et al. [9] considered a dictionary learning approach based on the difference in convex programming and further improved the DFL accuracy with a tracking neighborhood rule. Recently, Huang et al. [18] and Han et al. [19] proposed improved sparse coding algorithms (ISCAs) with log-regularizers, which positioned targets in more complex environments robustly. To overcome the low accuracy and insufficient robustness of DFL, Zhao et al. [20] considered a block sparse scheme, that achieved robust performance under severely noisy conditions. However, most of the above algorithms failed to consider the effect of high-dimensional data on computing efficiency. Recent work showed that sparse coding via the iterative shrinkage thresholding algorithm (SC-ISTA) [11] with the L1 Norm as penalty function has good localization accuracy and robustness. However, the L1 Norm applied in the above algorithm still has room for improvement in seeking a sparse solution to the L0 Norm.

Related research [19, 21–25] showed that using generalized penalty functions other than the L1 Norm to approximate the sparse solutions of the L0 Norm can achieve good performance with nonconvex minimization optimization. In a previous study [26], we proposed a sparse coding-iterative p-thresholding algorithm (SC-IpTA) and considered a thresholding function with parameter p as the penalty function to derive sparse solutions and achieve better DFL

performance, where the accuracy and robustness of the proposed method were verified with localization results.

To some extent, the accuracy is restricted by the number of sensor nodes. To achieve improved accuracy and enlarged areas of interest, most methods resort to more sensor nodes, leading to greater measurement data volume. With the RF model [27], the total number of wireless links between sensor node pairs increases dramatically with the number of sensor nodes, which places a greater burden on computing and data storage. In order to cater to the real-time requirements of IoT, the time cost of DFL systems needs to be reduced.

With high dimensionality, data mining is susceptible to the curse of dimensionality, a natural countermeasure of which is dimension reduction, where the high-dimensional feature space is projected onto a low-dimensional subspace. The most common dimension reduction method is principal component analysis (PCA) [28], which seeks a small number of orthogonal basis vectors to represent the maximum variance of the variables in the dataset.

The dimensionality of dictionaries can be reduced through the orthogonal transformation with PCA, which converts a series of possibly correlated vectors to linearly unrelated vectors, termed principal components. In order to reduce dimensionality, new features should be identified to reveal the main characteristics of the original data, each of which is a linear combination of orthogonal features. Eigenvalue decomposition (EVD), commonly utilized for signal features extraction and data representation, has wide applications such as image compression.

Recently, the research on DFL has shifted its focus to efficient computation after dimension reduction. To reduce the amount of localization data and the storage costs of DFL systems, Liu et al. [29] proposed a two-level controlling redundancy reduction approach for indoor DFL based on PCA node reduction. On the basis of [9], Li et al. [30] further proposed an outlier suppression approach via non-convex robust PCA for safe localization with dimension reduction. Wang et al. [31] achieved device-free simultaneous wireless localization and activity recognition (DFLAR) with a wavelet feature using RSS signal features extracted by PCA. Shi et al. [32] presented a computationally efficient approach for device-free indoor location tracking systems based on channel state information (CSI) metrics, which proved efficient and robust for high-dimensional CSI vectors due to the PCA-based dimension reduction. However, those methods mentioned above only considered one single target. Huang et al. [11] performed sparse coding with subspace techniques in low-dimensional signal subspace, achieving high localization accuracy and robustness while reducing time cost.

In this paper, SC-IpTA based on generalized thresholding is first introduced, which achieves good accuracy with high robustness. Then, an enhanced generalized thresholding scheme based on dimension reduction is derived to improve the computational and storage efficiencies of DFL systems, which overcomes the increased computation complexity induced by the high-dimensional RSS sampling data. In the meantime, the multi-target performance of the proposed method is also analyzed.

Firstly, the original RSS data for a single target and multiple targets were collected from the Sensing and Processing Across Networks (SPAN) Lab at the University of Utah [27]. Then, the overcomplete dictionary $\mathbf{D}$ was constructed with RSS data via matrix transformation. After that, dimension reduction was performed for each row and column of matrix $\mathbf{D}$ with PCA to yield the subspace dictionary matrix $\mathbf{D}^*$. In addition, the same operation is performed for the transformation from the observation vector b to $\mathbf{b}^*$, as depicted in Fig. 2. Next, the objective function constructed by the regularized constraint term with a p generalized thresholding was formulated as a non-convex optimization problem. Finally, the sparsest solution of the underdetermined equation was obtained using SC-IpTA, and the target location was estimated.

Contributions of the paper are summarized in the following three aspects. 1) A subspace sparse coding iterative p-thresholding algorithm(SSC-IpTA) is proposed by reducing the dimension of learning dictionaries and retaining the trunk of the most important information to provide a higher computing efficiency and save data storage space. 2) The robustness of the algorithm is evaluated numerically under different ambient noise levels. 3) The multi-target accuracy of the proposed algorithm is verified.

The structure of the paper is as follows. In Section 2, DFL systems and SRC models are introduced. Section 3 elaborates on the improved SSC-IpTA. In Section 4, the performance of the proposed algorithm is evaluated. In Section 5, conclusions are drawn.

2 DFL systems and the SRC model

2.1 DFL systems

In DFL systems, multiple sensor nodes are installed opposite each other on the edge of the areas of interest, as shown in Fig. 3. In order to facilitate the inference of the target location, the area is divided into grids, namely reference points (RPs), and the target is assumed to be in one of the RPs. Each sensor sends radio while the rest receive RSS signals simultaneously, forming sensor pairs.As the target moves into a different RP, the wireless links passing

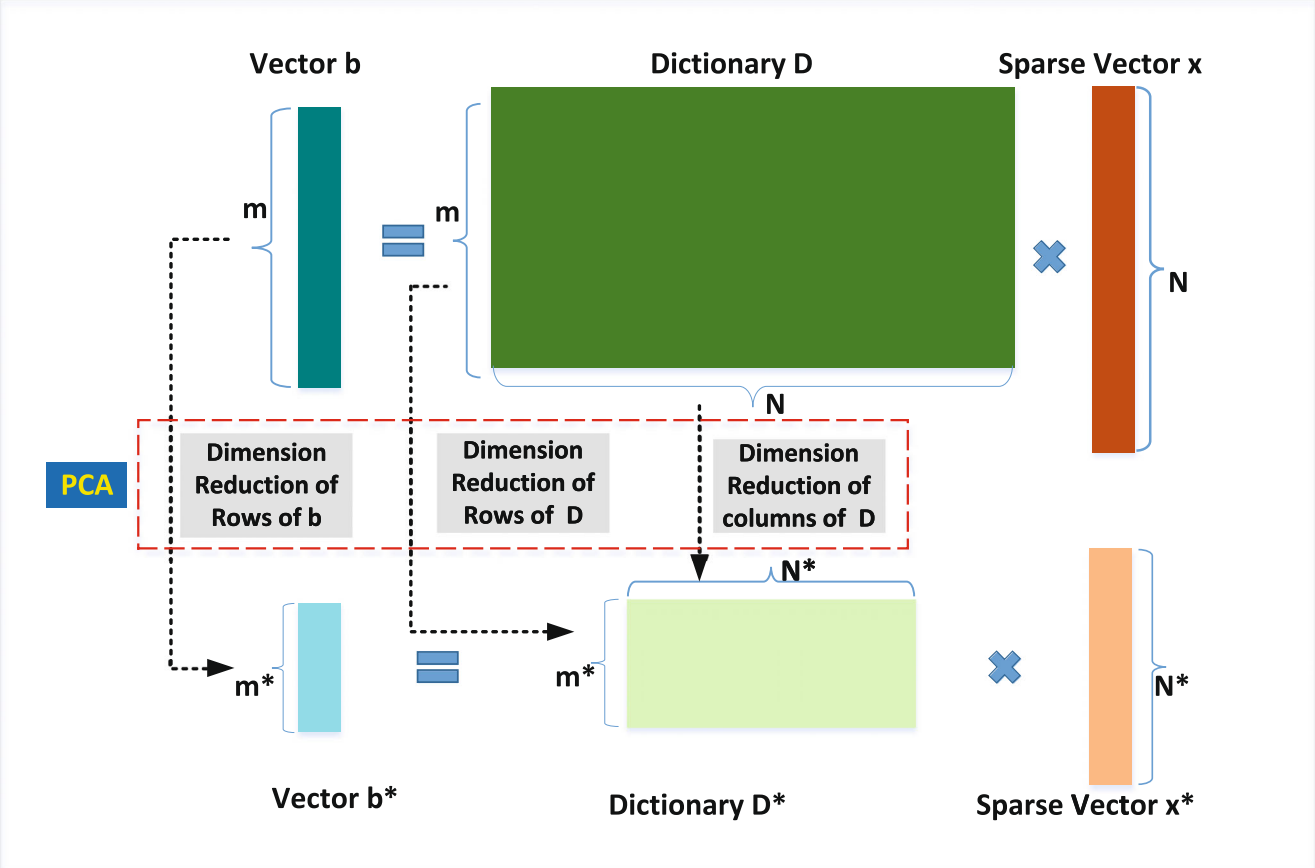

Fig. 2 The SRC model based on PCA

through it are blocked, and the received RSS measurement configurations differ, showing the specific characteristics of the target location. Thus the target location can be inferred by detecting the changes in RSS signals.

Since the above RSS measurement configurations represent different classes, they can be considered an image classification problem. In addition, the targets are far fewer than RPs in a typical DFL system.Hence the characteristics of sparsity. Therefore, the problem can be dealt with under SRC.

2.2 SRC model and sparse coding

SRC [33] is adopted to recover the observation signal from the underdetermined equation with overcomplete dictionaries with the constraints of the minimum reconstruction error. In addition, the observed vector **b** can be expressed as a linear combination of dictionary matrix **D** and sparse coefficient vector **x**, as shown in Fig. 2. The solutions of coefficient vector via sparse coding are featured with sparsity, i.e., most elements are zero. The DFL problem can be

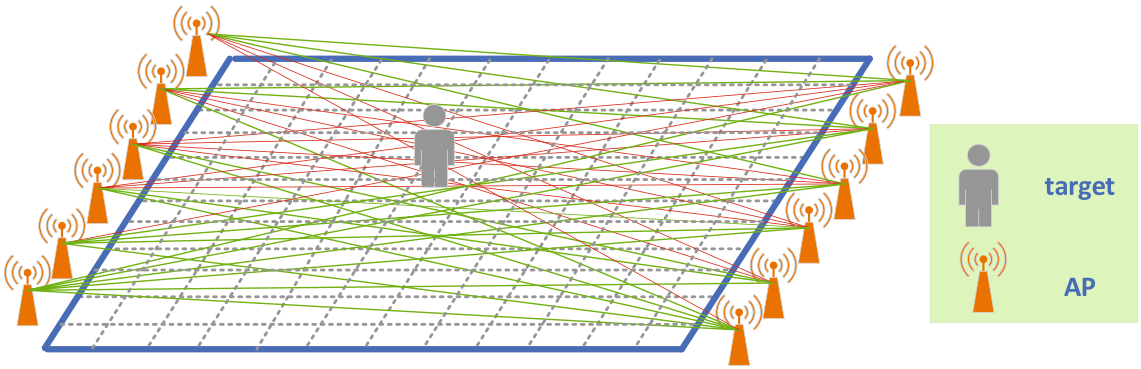

Fig. 3 A typical DFL system

solved by mapping the index of the sparse coefficient vector with the target location.

Sparse coding is the process of computing coefficient vectors based on a set of observed signals and known dictionaries, i.e., solving inverse problems.

The linear equation for Fig. 3 can be represented as:

$$\mathbf{b}^{m\times 1} = \mathbf{D}^{m\times N}\mathbf{x}^{N\times 1} \tag{1}$$

DFL systems require more RP measurement repetitions to construct the dictionary for better accuracy of the target localization in the offline stage, which causes N to become larger than m. When $N > m$, the equation becomes an underdetermined system with none-unique solutions, making it an ill-posed problem. With proper sparse constraints, the sparsest solution can be obtained, and the problem becomes well-posed.

Since the sparsest solution based on the L0 Norm is an NP-hard problem [34], suboptimal solutions are solved by orthogonal matched pursuit (OMP) algorithms, which do not apply to the DFL problem with high-dimensional data due to inefficiency. According to [13] and [14], the sparsest solution can be obtained by relaxing the L0 Norm to an L1 Norm optimization problem. The regularization method is used to solve combinatorial optimization problems consisting of an error term and an L1 penalty function term, i.e.:

$$\mathbf{x}^* = arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{b} - \mathbf{D}\mathbf{x}\|_2^2 + \tau\|\mathbf{x}\|_1 \tag{2}$$

where τ is the regularization parameter.

Assuming that there are K APs, and $\theta_{i,j}$ represent the RSS measurements received by the i-th AP from the j-th AP, vector $\boldsymbol{\theta}_i$ can be expressed as:

$$\boldsymbol{\theta}_i = [\theta_{i,1}, \theta_{i,2}, \cdots, \theta_{i,K}]^H \tag{3}$$

where $\theta_{i,i}$ is equivalent to the signal value transmitted by the i-th AP and H is the conjugate transpose operator. All $\boldsymbol{\theta}_i$ representing RSS values of the links compose matrix Θ as follows.

$$\Theta = [\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \cdots, \boldsymbol{\theta_K}]^H \tag{4}$$

The operation of the DFL system consists of two stages: offline training and online testing.

1) At the offline training stage, the area of interest of the DFL system is discretized into G RPs, each labeled with an RP index. Suppose the target is in the g-th RP and Z RSS matrices can be produced as follows:

$$\{\Theta_{g,1}, \Theta_{g,2}, \cdots, \Theta_{g,t}, \cdots, \Theta_{g,Z}\} \tag{5}$$

where $\Theta_{g,t}$ is the t-th measurement RSS matrix of the g-th RP, and Z is the repeated number of sampling at each RP.

The RSS matrix can be transformed into a column vector $\mathbf{d}_{g,t}$. Thus, the following matrix $\mathbf{D}_g$ can be constructed with the Z sample vector of the g-th AP as the column.

Repeating the experiment Z trials produces G sample matrices, and $N = G \times Z$ sample vectors are combined to construct a matrix $\mathbf{D}$.

2) At the online testing stage, the observed signal with the target in the area of interest is the vectorization of RSS matrix Θ .

The target assumed to be in the g-th RP, belongs to the g-th class. If sufficient sampling is taken in the g-th RP to ensure $N > m$, the observed signal vector $\mathbf{b}$ can be approximated using sampling matrix $\mathbf{D}_g$.

$$\mathbf{b} = \mathbf{D}_g\mathbf{x}_g = \sum_{t=1}^{Z} \mathbf{d}_{g,t}x_{g,t} \quad (0 \le g \le G) \tag{6}$$

In (6), $\mathbf{x}_g = [x_{g,1}, x_{g,2}, \cdots, x_{g,t}, \cdots, x_{g,Z}] \in \mathbf{R}^Z$ is the sparse coefficient vector, and $x_{g,t} \in \mathbf{R}$ is the coefficient of the element.

According to [11] and [26], the linear representation of the observation vector can be sparsely represented with N sample vectors of the dictionary.

$$\mathbf{b} = \mathbf{D}_g\mathbf{x}_g = \mathbf{D}\mathbf{x} \tag{7}$$

where

$$\begin{aligned}\mathbf{x} &= [0, \cdots, \mathbf{x}_g, \cdots, 0]^H \\ &= [0, \cdots, \underbrace{x_{g,1}, x_{g,2}, \cdots, x_{g,t}, \cdots, x_{g,Z}}_{\mathbf{x}_g}, \cdots, 0]^H\end{aligned} \tag{8}$$

Equation (7) is a sparse representation problem that can be solved by SRC, which reconstructs the input observed signal with a few atoms selected from the overcomplete dictionary and achieves the minimum reconstruction error.

2.3 Sparse Coding-Iterative p-Thresholding Algorithm

To solve the L1 minimization problem in (2), Huang et al. [11] presented the SC-ISTA using the soft thresholding operator.

$$(\mathbf{S}_\tau(\mathbf{x}))_k = \text{sgn}(\mathbf{x_k})\max\{0, \mid \mathbf{x_k} \mid -\tau\} \tag{9}$$

To solve the optimization problem, we introduced a generalized thresholding operator in [26], as can be defined as follows:

$$\mathbf{S}_\tau^p(\mathbf{x}_i) = \text{sgn}(\mathbf{x}_i)\max\{0, \mid \mathbf{x}_i \mid -\tau \mid \mathbf{x}_i \mid^{p-1}\} \tag{10}$$

The algorithm can be developed using the thresholding function as:

$$\mathbf{x}_{i+1} \leftarrow \mathbf{S}_\tau^p(\mathbf{x}_i + \mathbf{D}^T\mathbf{b} - \mathbf{D}^T\mathbf{D}\mathbf{x}_i) \tag{11}$$

By solving (10) and (11), the sparse solution of (2) can be obtained.

$$\mathbf{x}^* = \{\mathbf{x}_1, \cdots, \mathbf{x}_N\} = \{\mathbf{x}_{1,1}, \cdots, \mathbf{x}_{g,t}, \cdots, \mathbf{x}_{G,Z}\} \tag{12}$$

where $t = \{1, \cdots, Z\}$. With $\mathbf{x}_g^* = \sum_{t=1}^{Z} x_{g,t}^*$, (12) can be transformed as follows:

$$\mathbf{x}^* = \{\mathbf{x}_1^*, \cdots, \mathbf{x}_g^*, \cdots, \mathbf{x}_G^*\} \tag{13}$$

where $\mathbf{x}_i^* = \{x_{i,1}^*, \cdots, x_{i,Z}^*\}$.

The location of single target is estimated to be within the Φ-th RP, where Φ is the index of the maximal nonzero element determined as follows:

$$\Phi = arg \max_g \{\mathbf{x}_1^*, \cdots, \mathbf{x}_g^*, \cdots, \mathbf{x}_G^*\} \tag{14}$$

For multiple targets, we can estimate T targets at the Φ_1-th RP,..., Φ_T-th RP, where $\Phi_1, \cdots, \Phi_T$ are the elements with decreasing order in $\mathbf{x}^*$, given by:

$$\{\Phi_1, \cdots, \Phi_T\} = arg \max_{1,\cdots,T} \{\mathbf{x}_1^*, \cdots, \mathbf{x}_T^*, \cdots, \mathbf{x}_G^*\} \tag{15}$$

Indices $\{\Phi_1, \cdots, \Phi_T\}$ of the maximum nonzero elements of $\mathbf{x}^*$ can be considered the target locations.

3 Proposed algorithm

Previous research has shown that existing OMP, BP, and ISTA take more time as the data grow, which affects their application in IoT scenarios with high real-time requirements. To address the issue, a subspace SC-IpTA (SSC-IpTA) is presented. Firstly, the columns and rows of the dictionary matrix constructed by the RSS measurements were reduced. Next, a new dictionary was built in the principal component subspace, the dimension number of which was determined by the cumulative energy ratio (CER). Then, the coefficient vector was obtained via the sparse coding algorithm. Lastly, the experimental evaluation was conducted to verify the algorithm's performance.

The flowchart of the proposed SSC-IpTA is summarized in Fig. 4.

The dictionary should be constructed and normalized after the original RSS data are ready in the offline training stage.

3.1 Column dimension Reduction for dictionary D

Dictionary $\mathbf{D}$ consisting of the original RSS data can be expressed as:

$$\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \cdots, \mathbf{D}_g, \cdots, \mathbf{D}_G] \tag{16}$$

$$\mathbf{D}_g = [\mathbf{d}_{g,1}, \cdots, \mathbf{d}_{g,i}, \cdots, \mathbf{d}_{g,Z}] \in \mathbf{R}^{m \times Z} \quad (1 \le g \le G) \tag{17}$$

where $\mathbf{d}_{g,i}$ is the column vector of the i-th column of $\mathbf{D}_g$. The main steps of column dimension reduction are summarized as follows.

(1) Zero mean normalization for the dictionary

First, each column of $\mathbf{D}$ (denoted as $\mathbf{D}_i$) subtracts the mean of $\mathbf{D}$ to form a new column.

$$\mathbf{D}_i \leftarrow \mathbf{D}_i - \frac{1}{G} \sum_{i=1}^{G} \mathbf{D}_i \tag{18}$$

Then, the updated columns are divided by the standard variance, i.e.:

$$\mathbf{D}_i' \leftarrow \frac{\mathbf{D}_i - \frac{1}{G} \sum_{i=1}^{G} \mathbf{D}_i}{\sigma} \tag{19}$$

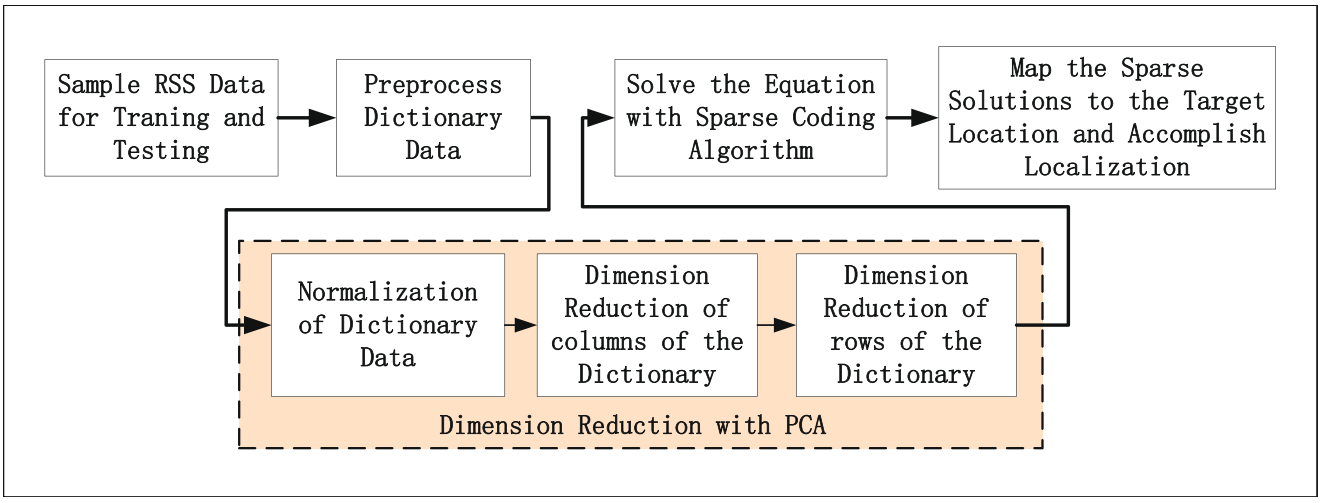

Fig. 4 Flowchart of the DFL system based on SSC-IpTA

Where the standard variance is defined as:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{G}(\mathbf{D}_i - \frac{1}{G}\sum_{i=1}^{G}\mathbf{D}_i)^2}{G-1}} \tag{20}$$

Thus the normalized matrix is:

$$\mathbf{D}_{nomal} = [\mathbf{D}'_1, \mathbf{D}'_2, \cdots, \mathbf{D}'_g, , \cdots, \mathbf{D}'_G] \tag{21}$$

Each column of the matrix $\mathbf{D}_{nomal}$ represents the different groups of the normalized value of the RSS matrix sampling.

(2) Calculation of the covariance matrix $\mathbf{C}$

At the g-th RP, we have:

$$\mathbf{D}'_g = [\mathbf{d}'_{g,1}, \mathbf{d}'_{g,2}, \cdots, \mathbf{d}'_{g,i}, \cdots, \mathbf{d}'_{g,Z}] \tag{22}$$

The covariance matrix $\mathbf{C}$ of $\mathbf{D}'_g$ in (24) is

$$\mathbf{C} = \mathbf{D}'_g \mathbf{D}'^H_g \tag{23}$$

or

$$\mathbf{C}_g = \frac{1}{Z}\sum_{j=1}^{Z}\mathbf{d}'_{g,j}\mathbf{d}'^H_{g,j} \tag{24}$$

where $\mathbf{C}_g \in \mathbf{R}^{m\times m}$.

(3) Singular value decomposition (SVD) on the covariance matrix

After SVD, we have:

$$\mathbf{C}_g = \mathbf{U}_g \sum\nolimits_g \mathbf{V}_g \tag{25}$$

where, $\mathbf{U}_g = \{\mathbf{u}_{g,1}, \mathbf{u}_{g,2}, \cdots, \mathbf{u}_{g,m}\}$, and $\mathbf{U}_g \in \mathbf{R}^{m\times m}$ is the normalized singular vector matrix whose vectors denote the most important features of dictionary $\mathbf{D}$. $\sum_g = diag\{\sigma_1, \cdots, \sigma_m\}$ is the singular value diagonal matrix, all singular values of which are sorted in descending order. The first column vector $\mathbf{u}_{g,1}$ is related to the maximum singular value σ_1, denoting the most important feature of matrix $\mathbf{D}_g$. As all RPs are transformed with SVD, all first column vectors of the matrix $\mathbf{U}_g$ of the overall RPs are selected to form a semi-transformation matrix $\mathbf{U}$ as follows:

$$\mathbf{U} = \{\mathbf{u}_{1,1}, \mathbf{u}_{2,1}, \cdots, \mathbf{u}_{g,1}, \cdots, \mathbf{u}_{G,1}\} \tag{26}$$

All vectors of matrix $\mathbf{U}$ correspond to the maximum singular values of the RPs.

3.2 Row dimension reduction for dictionary D and observed vector b

(1) Firstly, the above semi-transformation matrix $\mathbf{U}$ is processed with zero mean normalization. Then, EVD is performed to produce the covariance matrix Δ of the normalized matrix $\mathbf{U}_{nomal}$.

$$\Delta = \mathbf{U}_{nomal}\mathbf{D}^H_{nomal} = \mathbf{W}\sum\nolimits_\lambda \mathbf{W}^H \tag{27}$$

where $\mathbf{W}$ is the matrix constructed by the singular vector $\mathbf{w}_i$ of the matrix Δ, which can be expressed as follows:

$$\mathbf{W} = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_i, \cdots, \mathbf{w}_m\} \tag{28}$$

In (27), $\sum_\lambda$ is the diagonal matrix expressed as:

$$\begin{aligned}\sum\nolimits_\lambda &= diag\{\lambda_1, \lambda_2, \cdots, \lambda_i, \cdots, \lambda_m\},\\ &\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m\end{aligned} \tag{29}$$

where the element λ_i on the main diagonal is the eigenvalue in descending order. We can select the first k eigenvectors to form the projecting matrix $\mathbf{W}_k$ as follows:

$$\mathbf{W}_k = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k\} \tag{30}$$

(2) The selection of characteristic number k for the principal component space

Here, CER is defined as the ratio of all eigenvalues from 1 to k to the sum of all eigenvalues, i.e.:

$$CER = \frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{m}\lambda_i} \geq \chi, \quad (1 \leq k \leq m) \tag{31}$$

We should make a trade-off between dimension and localization precision by selecting the proper DFL CER higher than the threshold χ to meet the requirements. Here, $\chi = 99\%$ and the corresponding k were selected to guarantee the localization performance. As a result, the matrix $\mathbf{W}_k \in \mathbf{R}^{m\times k}$ can be determined as follows:

(3) The low-dimensional dictionary matrix $\mathbf{D}_k$ and the low-dimensional vector $\mathbf{b}_k$ can be obtained using the projecting matrix $\mathbf{W}_k$.

$$\mathbf{D}_k = \mathbf{W}^{\mathbf{H}}_{\mathbf{k}}\mathbf{U} \in \mathbf{R}^{k\times G} \tag{32}$$

$$\mathbf{b}_k = \mathbf{W}^{\mathbf{H}}_{\mathbf{k}}\mathbf{b} \in \mathbf{R}^{k\times 1} \tag{33}$$

Finally, low-dimensional sparse solutions can be obtained by solving the optimization problem in (34), which is similar to the procedure of SC-IpTA.

$$\mathbf{x}^* = arg\min_{\mathbf{x}}\frac{1}{2}\|\mathbf{b}_k - \mathbf{D}_k\mathbf{x}\|_2^2 + \tau\|\mathbf{x}\|_1 \tag{34}$$

The pseudocode of SSC-IpTA is as follows.

Algorithm 1 SSC-IpTA.

Require: Dictionary Matrix $\mathbf{D} \in \mathbf{R}^{m\times N}$;initial value $x_0 \in \mathbf{R}^N$; observed vector $\mathbf{b} \in \mathbf{R}^{m\times 1}$; regularization parameter $\tau < ||\mathbf{D}^H\mathbf{b}||_\infty^{2-p}$; threshold value ε;maximum number of iterations $\mathbf{M}$; shrinkage parameter $p \in \mathbf{R}$

Ensure: Sparse vector $\mathbf{x}$

1.column dimension reduction

1)Normalizing the dictionary $\mathbf{D}$ as $\mathbf{D}_{nomal}$ by (16)(17)

2)Calculating $\mathbf{C}$ by (23)(24)

3)Performing SVD on $\mathbf{C}_g = \mathbf{U}_g \sum_g \mathbf{V}_g$

2.row dimension reduction

1)Performing EVD on $\Delta = \mathbf{U}_{nomal}\mathbf{U}_{nomal}^H = \mathbf{W}\sum_\lambda \mathbf{W}^H$

2)Determing k while satisfying (31) $\mathbf{W}_k \leftarrow \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k\}$

3)Performing $\mathbf{D}_k = \mathbf{W_k^H U}$, $\mathbf{b}_k = \mathbf{W_k^H b}$

3.Performing $\mathbf{S}_\tau^p(\mathbf{x}_i) \leftarrow \text{sgn}(\mathbf{x}_i)\max\{0, \mid \mathbf{x}_i \mid -\tau \mid \mathbf{x}_i \mid^{p-1}\}$

for $i = 0$ to M **do**

repeat

compute $\mathbf{x}_{i+1} \leftarrow \mathbf{S}_\tau^p(\mathbf{x}_i + \mathbf{D}_k^H\mathbf{b}_k - \mathbf{D}_k^H\mathbf{D}_k\mathbf{x}_i)$;

until $\|\mathbf{x}_{i+1} - \mathbf{x}_i\| \le \varepsilon$

end for

$\mathbf{x} \leftarrow \mathbf{x}_{i+1}$

$\mathbf{x}$ is transformed by (13);

if locate one target then **then**

Let Φ be the Index of the target's location by (14)

end if

return Φ

if locate Multi-targets then **then**

Let $\{\Phi_1, \cdots, \Phi_T\}$ as the indices of multi-targets' locations by (15)

end if

return $\{\Phi_1, \cdots, \Phi_T\}$

4 Performance evaluation

4.1 Experimental setups

A dataset is dowloaded from the SPAN Lab [27]. The monitoring area is depicted in Fig. 5, which is 21 feet by 21 feet outdoor square divided into 35 RPs (the third grid for the bottom left is occupied by a tree). A total of 28 APs were installed on the parameters. A more detail configuration can be found in [11, 27]. In each grid, every 30 RSS measurement samples are gathered, of which 25 samples are used to construct the learning dictionary for training and the observation vector for testing, respectively. The experiments were performed in MATLAB R2019b and a 64-bit Windows 10 computer with 8 GB RAM and Intel E3-1230 CPU @ 3.3 GHz.

4.2 Experiment metrics and other settings

To evaluate the performance of the proposed method, alternative algorithms such as SSC-OMP and SSC-ISTA were also applied with L0 and L1 regularization terms, respectively, as references.

Localization accuracy was employed as the metric to compare the performance of the algorithms. The accuracy is defined as the ratio of correctly estimated samples N_c to the total test samples N_t.

Assuming The actual target location in the i-th RP (x_i, y_i) and the corresponding estimated location coordinate is $(\bar{x}_i, \bar{y}_i)$, the distance between the two sets of coordinates is the localization error (LE). The average localization error (ALE) is the average value LE at each RP.

$$ALE = \sum_{i=1}^{G} \frac{\sqrt{(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2}}{G} \tag{35}$$

Since the RSS samples are inevitably interfered with by environment noise, different levels of Gaussian noise were added to enhance the robustness.

Table 1 lists the typical metrics for the experiment.

4.3 Results and discussion

4.3.1 Performance analysis of single target localization in low-dimensional space

It is assumed that the target is in the 14th grid in Fig. 5, and the signal-to-noise ratios (SNRs) of the dictionary and the observed signal are both 20 dB.

Different levels of noise are added to the dictionary to evaluate the ALE of the algorithm at the observed signal SNR of 20 dB, as shown in Fig. 6. As k increases from 10 to 784, the ALE decreases from 12 feet to nearly zero. Especially when $k = 15$, the performance of SSC-IpTA (green line) is similar to that of SC-IpTA (pink line) without dimension reduction.

Table 1 Typical value of numerical computation

Parameter	Value
Maximum Iteration Number M	200
Regularization parameter τ	0.1
Thresholding value ε	0.1
Shrinkage parameter p	0.9
Typical characteristic number k	5,10,15,20,25,784

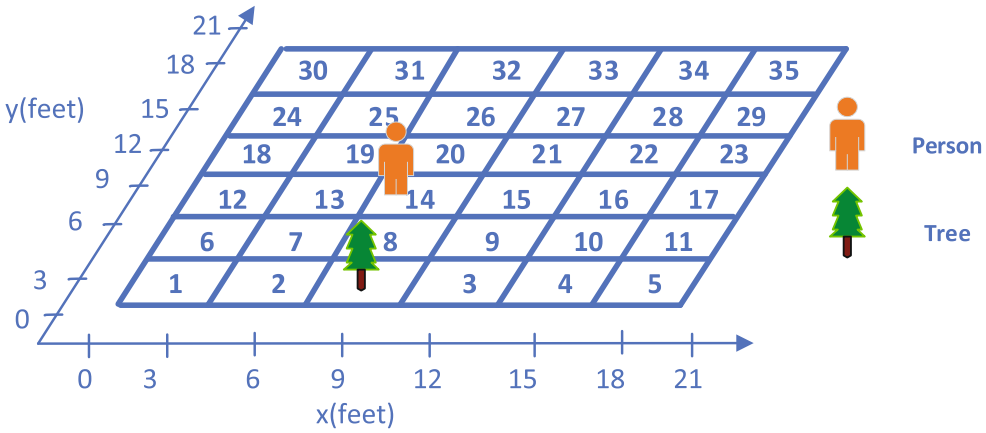

Fig. 5 Experiment diagram of DFL based on the dataset of the SPAN lab

As shown in Fig. 7, the accuracy of SSC-IpTA and the other algorithms improves as the characteristic number k increases. When k increases from 12 to 15, the accuracy of SSC-IpTA increases from 30% to almost 100%, which is superior to SSC-ISTA and comparable to SSC-OMP.

According to Fig. 8, ALE decreases to zero as the SNR of the dictionary increases under the condition that $k = 20$ and the dimension of the dictionary is 20×35. The ALE of SSC-IpTA decreases rapidly as the SNR of the observed signal increases from 15 dB to 20 dB and to noiseless. At dictionary SNRs above 15 dB, ALE is approximately zero when the observed signal SNRs are 20 dB and noiseless.

Table 2 shows the computation time of SC-OMP, SC-ISTA, and SC-IpTA at the RP index of 15 and the maximum number M of 200. The time cost of SC-ISTA is lower than those of SC-IpTA and SC-OMP before dimension reduction. In addition, the computing efficiencies of all algorithms are improved by two or three orders of magnitude with dimension reduction. Moreover, the efficiency of SSC-IpTA is superior to SSC-ISTA and SSC-OMP, thus satisfying the real-time requirements of IoT.

4.3.2 Comparison with state-of-the-art DFL methods

The proposed method was compared with another two novel state-of-the-art DFL algorithms, ISCA [18] and BSCPO [20],based on the same dataset.

According to the Fig. 6 in [26], the performance of SC-IpTA can be improved by measureing the sparsity with the distinctive capability of the proposed generalized thresholding algorithm with adaptive parameter p. Meanwhile, most of the useful information can be extracted with PCA. The comparison results are presented in Table 3. The proposed algorithm achieve the highest localization accuracy at the

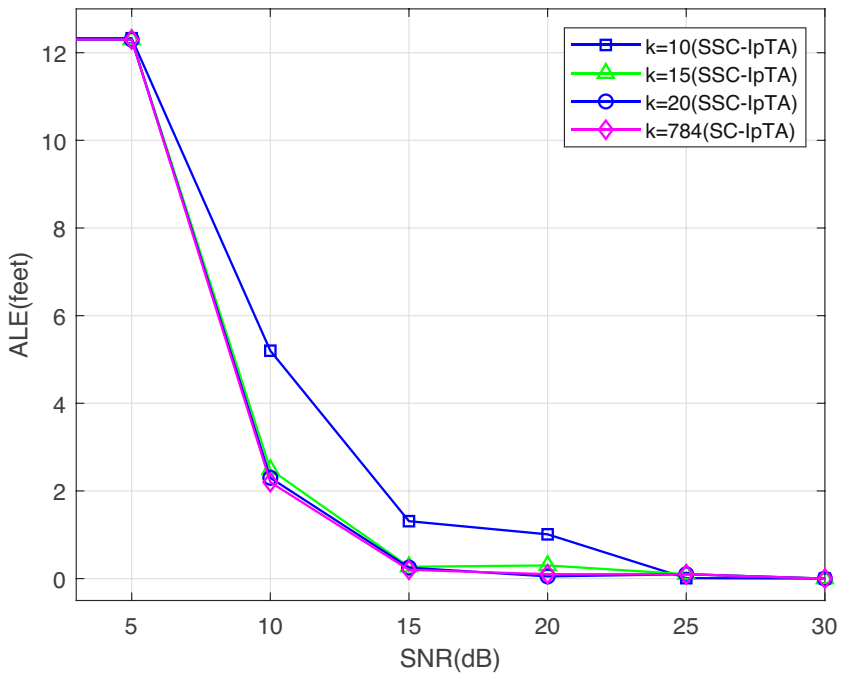

Fig. 6 Comparison of ALE of the algorithm with the characteristic number k of 10, 15, 20, and 784

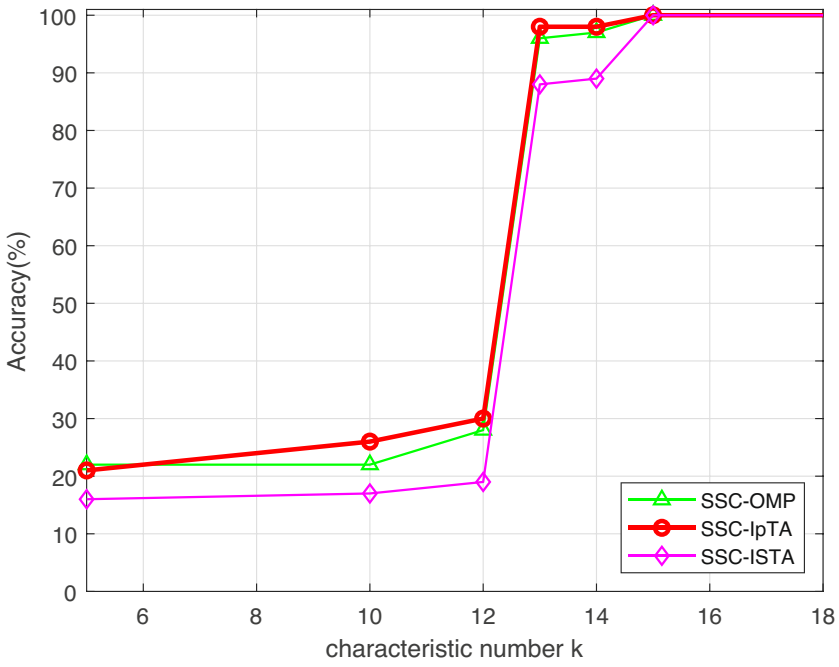

Fig. 7 Comparison of accuracy between the SSC-IpTA and alternative algorithms with different characteristic numbers k

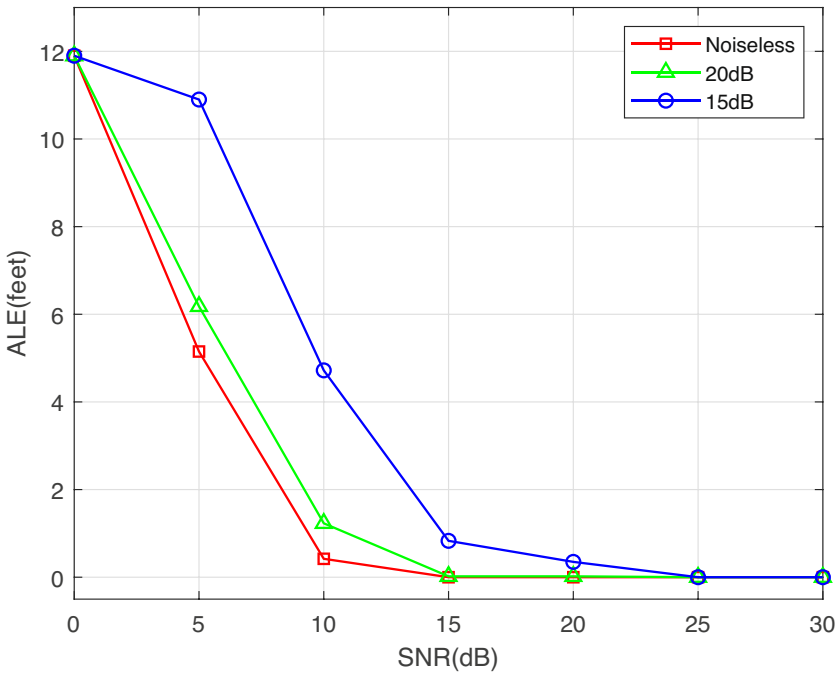

Fig. 8 Comparison of ALE between the algorithms with the observed signal SNR of noiseless, 20 dB, and 10 dB, $k = 20$, and the dictionary dimension of 20×35

Table 2 Time costs of different algorithms

Algorithm	Time cost(s)	Algorithm	Time cost(s)
SC-OMP	0.4011	SSC-OMP	0.000339
SC-ISTA	0.0104	SSC-ISTA	0.000398
SC-IpTA	0.0241	SSC-IpTA	0.000151

Table 3 Performance comparison with other algorithms

Compared Algorithm	SNR=15dB(dictionary)
SSC-OMP [10]	71.2%
SSC-ISTA [11]	85.4%
SSC-ISCA [18]	85.0%
SSC-BSCPO [20]	86.1%
SC-IpTA [26]	90.6%
The proposed	100%

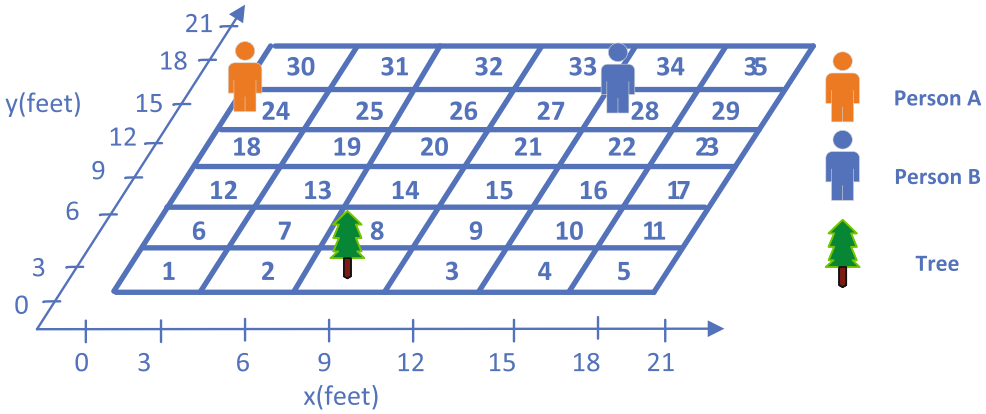

Fig. 9 Illustration of the DFL system with multiple persons ($N = 2$)

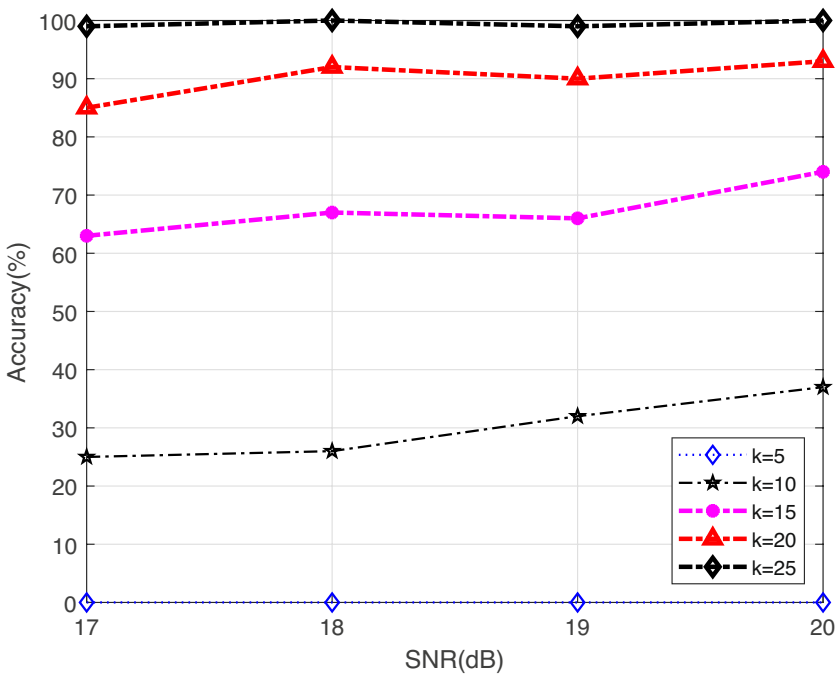

Fig. 10 Accuracy of SSC-IpTA with different k values for multiple targets ($N = 2$)

dictionary data SNR of 15dB, the test signal SNR of 20 dB, and k=15, i.e., the proposed method outperforms the other algorithms in terms of robustness and accuracy.

4.3.3 Performance analysis of multi-targets localization

It is assumed that $N(N > 1)$ targets, e.g., $N = 2$, are in the area of interest, one at RP-24 and the other at RP-28, as depicted in Fig. 9. As shown in Fig. 10, the accuracy levels of SSC-IpTA are compared under different k values and the dictionary SNRs of 17 dB, 18 dB, 19 dB, and 20 dB, respectively. the results show that when k is below 5, the accuracy is nearly zero. In contrast, when k is above 15, the accuracy reaches 60%. When $k = 25$, the performance is similar to that when $k = 784$, and the accuracy is almost up to 100%.

5 Conclusion

In order to improve the computation efficiency and reduce the time delay of DFL system while maintaining high accuracy, a generalized thresholding algorithm based on dimension reduction with PCA for DFL was proposed in this paper. The method formulates the sparse model based on the DFL problem into a subspace problem. Numerical results showed that the proposed algorithm could improve the computation efficiency of DFL systems and take significantly less time than other algorithms, implying its applicability to IoT scenarios with high real-time requirements.

Acknowledgments This work is supported in part by the National Natural Science Foundation of China under Grant 61771258, the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grants KYCX20_0732, KYCX21_0749, and KYCX20_0739, and the Open Research Fund of Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications), Ministry of Education.

Data Availability The authors declare that all data generated or analysed during this study are included in this article.

Declarations

Competing interests There are no competing interests to declare.

References

1. Pahlavan K, Krishnamurthy P, Geng Y (2015) Localization challenges for the emergence of the smart world. IEEE Access 3:3058–3067. https://doi.org/10.1109/ACCESS.2015.2508648
2. Youssef M, Mah M, Agrawala A (2007) Challenges: device-free passive localization for wireless environments. In: Proceedings of the 13th annual ACM international conference on mobile computing and networking, pp 222–229
3. Zhang D, Ma J, Chen Q, Ni LM (2007) An rf-based system for tracking transceiver-free objects. In: Fifth annual IEEE international conference on pervasive computing and communications (PerCom'07), pp 135–144. IEEE
4. Wilson J, Patwari N, Vasquez FG (2009) Regularization methods for radio tomographic imaging. In: 2009 Virginia tech symposium on wireless personal communications
5. Zhao L, Huang H, Li X, Ding S (2019) An accurate and robust approach of device-free localization with convolutional autoencoder. IEEE Internet of Things Journal 6(3):5825–5840. https://doi.org/10.1109/JIOT.2019.2907580
6. Zhao L, Su C, Huang H, Han Z, Ding S, Li X (2019) Intrusion detection based on device-free localization in the era of iot. Symmetry 11(5):630. https://doi.org/10.3390/SYM11050630
7. Zhang Z, Xu Y, Yang J, Li X, Zhang D (2015) A survey of sparse representation: algorithms and applications. IEEE Access 3:490–530. https://doi.org/10.1109/ACCESS.2015.2430359
8. Wang DS, Guo XS, Zou YX (2016) Accurate and robust device-free localization approach via sparse representation in presence of noise and outliers. In: 2016 IEEE International conference on digital signal processing (DSP), pp 199–203
9. Li X, Ding S, Li Z, Tan B (2017) Device-free localization via dictionary learning with difference of convex programming. IEEE Sensors J 17(17):5599–5608. https://doi.org/10.1109/JSEN.2017.2730226
10. Liu T, Luo X, Liang Z (2018) Enhanced sparse representation-based device-free localization with radio tomography networks. J Sens Actuator Netw 7(1):7. https://doi.org/10.3390/jsan7010007
11. Huang H, Zhao H, Li X, Ding S, Zhao L, Li Z (2018) An accurate and efficient device-free localization approach based on sparse coding in subspace. IEEE Access 6:61782–61799. https://doi.org/10.1109/ACCESS.2018.2876034
12. Huang H, Han Z, Ding S, Su C, Zhao L (2019) Improved sparse coding algorithm with device-free localization technique for intrusion detection and monitoring. Symmetry 11:637. https://doi.org/10.3390/SYM11050637
13. Tropp JA, Gilbert AC (2007) Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inform Theory 53(12):4655–4666. https://doi.org/10.1109/TIT.2007.909108
14. J. CE (2006) Compressive sampling. In: Proceedings of the international congress of mathematicians, vol 3, pp 1433–1452
15. Daubechies I, Defrise M, De Mol C (2004) An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun Pur Appl Math 57(11):1413–1457. https://doi.org/10.1002/CPA.20042
16. Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imag Sci 2:183–202. https://doi.org/10.1137/080716542
17. Selesnick IW (2009) Sparse signal restoration. Connexions, 1–13
18. Huang H, Zhang C, Wu H, Dai Z, Zhao L, Su C (2021) An improving sparse coding algorithm for wireless passive target positioning. Phys Commun 49:1–9. https://doi.org/10.1016/j.phycom.2021.101487
19. Han Z, Su C, Ding S, Huang H, Zhao L (2019) Device-free localization via sparse coding with log-regularizer. In: 2019 IEEE 10th international conference on awareness science and technology (iCAST), pp 1–6
20. Zhao L, Huang H, Su C, Ding S, Huang H, Tan Z, Li Z (2021) Block-sparse coding-based machine learning approach for dependable device-free localization in iot environment. IEEE Internet Things J 8:3211–3223. https://doi.org/10.1109/JIOT.2020.3019732
21. Woodworth J, Chartrand R (2016) Compressed sensing recovery via nonconvex shrinkage penalties. Inverse Probl 32(7):075004. https://doi.org/10.1088/0266-5611/32/7/075004
22. Chartrand R (2009) Fast algorithms for nonconvex compressive sensing: Mri reconstruction from very few data. In: 2009 IEEE international symposium on biomedical imaging: from nano to macro, pp 262–265

23. Chartrand R (2007) Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Process Lett 14(10):707–710. https://doi.org/10.1109/LSP.2007.898300
24. Chartrand R, Staneva V (2008) Restricted isometry properties and nonconvex compressive sensing. Inverse Probl 24:035020. https://doi.org/10.1088/0266-5611/24/3/035020
25. Voronin S, Chartrand R (2013) A new generalized thresholding algorithm for inverse problems with sparsity constraints. In: 2013 IEEE international conference on acoustics, speech and signal processing, pp 1636–1640
26. Cheng Q, Zhang L, Xue B, Shu F, Yu Y (2021) Device-free localization via sparse coding with a generalized thresholding algorithm. IEICE Trans Commun E105-B:1. https://doi.org/10.1587/transcom.2021ebp3048
27. Wilson J, Patwari N (2010) Radio tomographic imaging with wireless networks. IEEE Trans Mob Comput 9(5):621–632. https://doi.org/10.1109/TMC.2009.174
28. Jolliffe I (2002) Principal component analysis. Springer, New York
29. Liu J, An H, MT PM, Cui Z, Zhao S (2017) Redundancy reduction for indoor device-free localization. Pers Ubiquit Comput 21(1):5–15. https://doi.org/10.1007/s00779-016-0979-8
30. Li X, Ding S, Li Y (2017) Outlier suppression via non-convex robust pca for efficient localization in wireless sensor networks. IEEE Sensors J 17(21):7053–7063. https://doi.org/10.1109/JSEN.2017.2754502
31. Wang J, Zhang X, Gao Q, Ma X, Feng X, Wang H (2016) Device-free simultaneous wireless localization and activity recognition with wavelet feature. IEEE Trans Veh Technol 66(2):1659–1669. https://doi.org/10.1109/TVT.2016.2555986
32. Shi S, Sigg S, Chen L, Ji Y (2018) Accurate location tracking from csi-based passive device-free probabilistic fingerprinting. IEEE Trans Veh Technol 67(6):5217–5230. https://doi.org/10.1109/TVT.2018.2810307
33. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2008) Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell 31(2):210–227. https://doi.org/10.1109/TPAMI.2008.79
34. Cevher V, Duarte MF, Baraniuk RG (2008) Distributed target localization via spatial sparsity. In: 2008 16th European signal processing conference, pp 1–5

Qin Cheng received his B.S. degree in electronic engineering from Changzhou Teacher's University of Technology, Changzhou, China, in 2002 and M.S. degree in electronic and communication engineering from Nanjing university of science and technology, Nanjing, China, in 2008. Since 2007, he has served as a lecturer at the Department of Electric Information Engineering, Jiangsu University of Technology, Changzhou, China. He is currently a PhD candidate in the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include machine learning and sparse signal processing in wireless sensor network.

Linghua Zhang received her M.S. degree in signal and information processing from Southeast University and her Ph.D. degree in signal and information processing from Nanjing University of Posts and Telecommunications, Nanjing, China, in 1990 and 2005, respectively. She is currently a professor and Ph.D. supervisor in the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China. Her research interests include speech signal processing, voice communication, and signal processing in wireless communication.

Bo Xue received his B.S. degree in electronic information engineering and his M.S. degree in signal and information processing from Yangzhou University, Yangzhou, China, in 2003 and 2006, respectively; and his Ph.D. degree in signal and information processing from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2018. He was an academic visitor from 2016 to 2017 with the Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada. Since 2018, he has served as an associate professor at the Department of Electric Information Engineering, Jiangsu University of Technology, Changzhou, China. His research interests focus on statistical signal processing and machine learning for wireless sensor networks.

Feng Shu received his bachelor's degree in telecommunication engineering from Southeast from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2012. He is currently a PhD candidate in the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include compressed sensing, deep learning and signal processing in wireless sensor network.

Affiliations

Qin Cheng[1,2] · Linghua Zhang[1] · Bo Xue[2,3] · Feng Shu[1] · Xu Wang[1]

Qin Cheng
cq@jsut.edu.cn

Bo Xue
yzuxb2003@163.com

Feng Shu
2018010101@njupt.edu.cn

Xu Wang
wangxhhtc@163.com

1 College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, 210003, Jiangsu, China

2 College of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, 213001, Jiangsu, China

3 Key lab of Broadband wireless communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Nanjing, 210003, Jiangsu, China