Soft Computing
https://doi.org/10.1007/s00500-018-3171-4

FOCUS

FSELM: fusion semi-supervised extreme learning machine for indoor localization with Wi-Fi and Bluetooth fingerprints

Xinlong Jiang[1,2,3] · Yiqiang Chen[1,2,3] · Junfa Liu[1,2,3] · Yang Gu[1,2,3] · Lisha Hu[4]

Abstract
Recently, the problem of indoor localization based on WLAN signals is attracting increasing attention due to the development of mobile devices and the widespread construction of networks. However, no definitive solution for achieving a low-cost and accurate positioning system has been found. In most traditional approaches, solving the indoor localization problem requires the availability of a large number of labeled training samples, the collection of which requires considerable manual effort. Previous research has not provided a means of simultaneously reducing human calibration effort and improving location accuracy. This paper introduces fusion semi-supervised extreme learning machine (FSELM), a novel semi-supervised learning algorithm based on the fusion of information from Wi-Fi and Bluetooth Low Energy (BLE) signals. Unlike previous semi-supervised methods, which consider multiple signals individually, FSELM fuses multiple signals into a unified model. When applied to sparsely calibrated localization problems, our proposed method is advantageous in three respects. First, it can dramatically reduce the human calibration effort required when using a semi-supervised learning framework. Second, it utilizes fused Wi-Fi and BLE fingerprints to markedly improve the location accuracy. Third, it inherits the beneficial properties of ELMs with regard to training and testing speeds because the input weights and biases of hidden nodes can be generated randomly. As demonstrated by experimental results obtained on practical indoor localization datasets, FSELM possesses a better semi-supervised manifold learning ability and achieves higher location accuracy than several previous batch supervised learning approaches (ELM, BP and SVM) and semi-supervised learning approaches (SELM, S-RVFL and FS-RVFL). Moreover, FSELM needs less training and testing time, making it easier to apply in practice. We conclude through experiments that FSELM yields good results when applied to a multi-signal-based semi-supervised learning problem. The contributions of this paper can be summarized as follows: First, the findings indicate that effective multi-data fusion can be achieved not only through data-layer fusion, feature-layer fusion and decision-layer fusion but also through the fusion of constraints within a model. Second, for semi-supervised learning problems, it is necessary to combine the advantages of different types of data by optimizing the model's parameters.

Keywords Fusion semi-supervised learning · Extreme Learning Machine · Indoor localization · Wi-Fi and Bluetooth fingerprints

1 Introduction

In recent years, indoor localization technology has undergone rapid development and attracted considerable interest both from the research community and industrial applications. Indoor localization can be defined as the task of inferring the location of a user by leveraging the received signal strength (RSS) information from multiple wireless sensor beacons. Over the past decade, many indoor localization methods based on wireless local area network (WLAN) signals (for example, Wi-Fi and Bluetooth), infrared signals (Lee et al. 2006) and ultrasound devices (González et al.

Communicated by X. Wang, M. Pelillo, A. K. Sangaiah.

✉ Yiqiang Chen
yqchen@ict.ac.cn

1 Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

2 Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing, China

3 University of Chinese Academy of Sciences, Beijing, China

4 Institute of Information Technology, Hebei University of Economics and Business, Shijiazhuang, Hebei, China

Published online: 06 April 2018

2009) have been presented. Among the existing approaches, WLAN-based approaches are regarded as potentially ideal methods because of their suitable effective distances and well-established infrastructures. At present, the most popular RSS-based location estimation techniques are fingerprint approaches, which consist of an offline training phase and an online positioning phase (Torres-Solis et al. 2010). During the training phase, location-annotated wireless signal fingerprints are collected to form a fingerprint database (Nguyen et al. 2005; Letchner et al. 2005). Then, in the positioning phase, a mobile device records its wireless signal fingerprint and compares it against the available fingerprints in the database.

To ensure high prediction accuracy, most machine learning methods (Gu and Sheng 2017; Bin 2015; Gu et al. 2015) require a large number of labeled samples (Gu et al. 2015; Wen et al. 2015; Gao et al. 2014; Chen et al. 2015). However, calibration is an enormous task that requires physical movement and manual labeling of the positions of interest (PoI), which represents a considerable barrier to practical application. To achieve high location accuracy, much previous research has attempted to take advantage of the fusion of data gathered from multiple wireless technologies, such as Wi-Fi and Bluetooth Low Energy (BLE), because Wi-Fi and BLE are two typical classes of WLAN signals for which the basic infrastructure has already been established in shopping malls, schools, hospitals, airports and other facilities. Comparatively speaking, existing single-signal-based methods (Bahl and Padmanabhan 2000; Haeberlen et al. 2004) cannot achieve high location accuracy due to the limited feature information available. Currently, most mobile devices (e.g., mobile phones) are equipped with abundant wireless technologies [e.g., Wi-Fi, BLE and near-field communication (NFC)], thus offering potential opportunities for indoor localization techniques based on data fusion (Rodrigues et al. 2012).

Regarding the first challenge mentioned above, only a few studies have sought to reduce the human effort required for calibration. Some authors Yang et al. (2012) and Zhou et al. (2014) have exploited users' motions and activities to assist in online location tracking and in obtaining a final location estimate. In this way, a location accuracy competitive with that of fingerprint calibration can be achieved. Other researchers have proposed a hybrid method (Ouyang et al. 2012) combining a probabilistic model with the Naïve Bayes (NB) and expectation maximization (EM) methods. Recently, effective methods of reducing human calibration effort have been developed based on the use of both labeled and unlabeled data. Both a hidden Markov model (HMM)-based tracking algorithm (Chai and Yang 2007) and a crowdsourcing approach (Yang et al. 2015) have been proposed to study the use of unlabeled data to improve location accuracy. Moreover, various semi-supervised learning methods using both labeled and unlabeled data have been attracting increasing attention (Chai and Yang 2005; Pan et al. 2007, 2006). In these methods, a graph Laplacian regularization of high-dimensional signal features is built and then mapped to physical locations in space. However, these methods still require a large amount of labeled data to ensure the location accuracy.

Regarding the second challenge, much work has been done to further enhance the location accuracy through data fusion. One effective way is to use Bluetooth signals to narrow down the potential zone in which an object should be located and then to utilize Wi-Fi signals to further estimate the final location (Aparicio et al. 2008). Other fusion methods simply combine RSS measures for different types of WLAN signals (Galvan-Tejada et al. 2012) or combine the location results corresponding to each individual signal (Galván-Tejada et al. 2013). Although these methods can certainly improve the accuracy of the results, they do not properly exploit the fundamental advantage of data fusion. For example, the effective distance for Wi-Fi signals is approximately 100 meters, whereas BLE has a much shorter effective distance of only 10 meters. Therefore, Wi-Fi-based methods perform better in large-scale localization scenarios because each Wi-Fi access point (AP) covers a large area, whereas BLE demonstrates better performance in small-scale applications such as room-level localization. The question of how to effectively combine Wi-Fi and BLE to achieve a good localization effect at both large and small scales remains an open issue.

Based on the above considerations, a problem that requires further study is how to achieve high location accuracy with little human calibration effort. In this paper, we propose an algorithm called fusion semi-supervised extreme learning machine (FSELM) to address this problem. The method presented here is an extension of our previous SELM approach (Liu et al. 2011). Unlike single-signal-based methods, our proposed approach comprehensively explores the information available from the fusion of Wi-Fi and BLE signals to achieve enhanced location accuracy. Moreover, FSELM uses a manifold regularization framework to realize a semi-supervised approach with reduced manual calibration effort. FSELM also inherits the beneficial properties of ELMs in terms of their fast training and testing speeds, which are achieved by adopting the random method of network parameter generation used in some previous works, such as feedforward neural networks with random weights (Schmidt et al. 1992) and random vector functional link (RVFL) networks without direct input–output links (Pao et al. 1994). The ELM algorithm also adopts the generalized least-squares (GLS) approach for obtaining an optimal solution, which is the same as that adopted in multivariable functional interpolation for adaptive networks (Lowe 1988). Experimental results indicate that FSELM can achieve better

localization performance with less calibration effort compared with some existing batch learning approaches [ELM (Huang et al. 2006), back propagation (BP) and support vector machine (SVM)] and semi-supervised learning methods (SELM, semi-supervised RVFL and fusion semi-supervised RVFL).

The remainder of this paper is organized as follows. Section 2 summarizes some related works on semi-supervised learning approaches and fused-signal-based indoor localization methods. Section 3 briefly reviews single-hidden-layer feedforward networks (SLFNs), ELMs and semi-supervised manifold regularization. Subsequently, Sect. 4 introduces the proposed FSELM methodology and procedure in detail. Then, Sect. 5 presents comparative results from experiments conducted in a shopping mall. Finally, we conclude the paper and discuss some future extensions of our method in Sect. 6.

2 Related work

The underlying idea of indoor localization is to use the RSS information measured from APs in a predefined wireless environment to estimate a user's current position via fingerprint matching or machine learning algorithms. RSS fingerprint-based methods of this kind have two phases: an offline training phase and an online positioning phase. During the offline phase, RSS information is collected from Wi-Fi or Bluetooth APs to build a mapping function between the RSS fingerprints and their corresponding coordinates. In the online phase, localization results can be obtained by inputting measured RSS information into this mapping function.

2.1 Semi-supervised approaches for indoor localization

In recent years, machine learning methods have proven to be an effective means of building mapping functions (Liu et al. 2007). However, a good prediction model must be based on an enormous amount of labeled training data. Generally speaking, the process of collecting such data is very time-consuming and tedious (Chen et al. 2005). By comparison, unlabeled data are much easier and cheaper to obtain. Thus, the question of how to train a prediction model using a large amount of unlabeled data combined with relatively little labeled data, known as sparsely calibrated location estimation, has become an important issue.

In previous approaches, manifold learning and semi-supervised learning methods have been widely adopted to solve this problem (Pan et al. 2007; Liu et al. 2011; Yang et al. 2015; Zhang and Zhi 2010; Pan et al. 2012; Lin et al. 2011; Chen et al. 2011; Scardapane et al. 2016; Xiang et al. 2015). Pan et al. (2007) proposed an online co-localization method of inferring the locations of users and APs. This method first uses a singular value decomposition (SVD)-based technique to obtain relative positions from labeled and unlabeled samples via dimensional reduction and then applies manifold regularization, with the help of labeled samples, to obtain the final localization result. Liu et al. (2011) developed a semi-supervised learning method named SELM (semi-supervised ELM) by introducing graph Laplacian regularization into the ELM algorithm. This method can achieve good performance with a fast learning speed on sparsely calibrated localization problems. Yang et al. (2015) adopted a crowdsourcing approach to gather large amounts of labeled and unlabeled data. Using both of labeled and unlabeled data, these authors took advantage of deep learning and semi-supervised learning to achieve a high location accuracy. Zhang and Zhi (2010) presented a semi-supervised mapping from signal space to physical coordinate space named LocMR. By computing the graph Laplacian regularization of the high-dimensional signal space, these authors accounted for both temporal and spatial constraints in an innovative way. To reduce calibration costs, Lin et al. (2011) represented the obtained samples in eigenvector space using the spectral decomposition of the Laplacian matrix and then applied an alignment strategy to label the unlabeled data. The experimental results showed that this algorithm can label unlabeled data with a high accuracy ($\sim$80%) based on only a small amount of labeled data ($\sim$20%), thereby effectively reducing the cost of data collection. Chen et al. (2011) proposed a semi-supervised Laplacian-regularized least-squares algorithm for localization in a wireless sensor network. This algorithm first uses an alignment criterion to learn an appropriate kernel function based on the similarities between anchors, and this kernel function is then used to measure the pairwise similarities between sensor nodes in the network. Subsequently, a semi-supervised learning algorithm based on manifold regularization is applied to obtain the locations of non-anchors.

However, all of the methods introduced above were proposed for solving single-signal-based localization problems. In the case of a multi-signal-based localization problem, they are not able to fully exploit the information provided by each signal. Scardapane et al. (2016) proposed a semi-supervised RVFL network based on a transductive framework. However, this approach is mainly intended to address binary classification problems and is not suitable for solving multi-class classification and regression problems.

2.2 Indoor localization based on signal fusion

The findings of the related works discussed above indicate that the location accuracy depends on not only the number of labeled samples but also the abundance of feature information obtained. Consequently, one effective means of improving the location accuracy is to consider multiple wireless signals.

Currently, most multi-signal-based localization technologies employ Wi-Fi and Bluetooth (Galvan-Tejada et al. 2012; Galván-Tejada et al. 2013; Aparicio et al. 2008; Cooper et al. 2016; Pandya et al. 2003; Hossain et al. 2007; Aparicio et al. 2009) because of the suitable effective distances and baffle-free nature of these technologies as well as their well-established infrastructures. Most multi-signal-based algorithms can be divided into two types, which we call combination algorithms and fusion algorithms.

In combination algorithms, location estimation is performed independently using every individual signal; then, the results are combined to obtain the final estimated position. In Pandya et al. (2003), Wi-Fi and Bluetooth signals are first separately considered in a RADAR model to obtain localization results, and the mean of the two results is then used as the final location estimate. Galván-Tejada et al. (2013) proposed a propagation model-based location method using Wi-Fi and BLE signals. These authors first separately computed the distances using a self-built BLE propagation model and a known Wi-Fi propagation model. Then, they presented an algorithm for obtaining the position of the user by combing the BLE- and Wi-Fi-based results. Aparicio et al. (2008) presented a signal combination method for Bluetooth and Wi-Fi, in which the localization process is divided into coarse-grained and fine-grained localization phases. In the first phase, the Bluetooth signal is used to narrow the potential location results to a small zone. Then, in the second phase, Wi-Fi signal measurements are used to locate the user only within the previously identified zone. This method offers significantly improved location accuracy at a reduced computation cost. Compared with the case of Wi-Fi alone, this multi-signal-based method can achieve higher location accuracy, with a mean error distance of less than 40 cm.

Unlike the methods above that use Wi-Fi and Bluetooth signals individually, algorithms of the second type are based on the construction of a fused map using Wi-Fi and Bluetooth power values. In Galvan-Tejada et al. (2012), a localization method based on a trilateration technique was presented. This article reported the achievement of propagation model based localization via the fusion of Wi-Fi and Bluetooth. First, the authors built propagation equations for both Bluetooth and Wi-Fi. Then, given the positions of both Bluetooth and Wi-Fi APs, the target position could be determined via the GLS method under the known constraints on the distances from the target to each AP. And in order to fuse the information obtained from Wi-Fi and BLE signals in the feature extraction layer. Cooper et al. (2016) proposed a method of fusing the Wi-Fi and BLE information obtained by most contemporary mobile devices and incorporated a boosting classification technique to achieve a significant improvement in terms of time consumption. However, fusion of this kind can often yield only limited improvement, and sometimes it may even reduce the location accuracy.

Therefore, in this paper, we pay closer attention to investigating how to effectively integrate the information obtained from multiple signals to improve the location accuracy.

3 Brief review of SLFNs, ELMs and semi-supervised manifold regularization

Before we can introduce our fusion-based semi-supervised neural network for multi-signal-based localization, we must briefly review single-hidden-layer feedforward networks (SLFNs), extreme learning machines (ELMs) Huang et al. (2006) and semi-supervised manifold regularization.

3.1 Single-hidden-layer feedforward networks (SLFNs)

An SLFN is a feedforward network with hidden nodes (radial basis function (RBF) or additive nodes) that can be described as follows:

$$f_M(x) = \sum_{i=1}^{M} \boldsymbol{\beta}_i G(\mathbf{x}; \boldsymbol{a}_i, b_i), \mathbf{x} \in R^n, \boldsymbol{a}_i \in R^n \tag{1}$$

where $\boldsymbol{a}_i$ and b_i denote the learning parameters related to the hidden nodes, $\boldsymbol{\beta}_i$ denotes the output weight connecting the ith hidden node to the output node, and $G(\mathbf{x}; \boldsymbol{a}_i, b_i)$ denotes the output of the ith hidden node related to the input $\mathbf{x}$. For additive hidden nodes using a threshold activation function or the sigmoid function, $g(x) : G(\mathbf{x}; \boldsymbol{a}_i, b_i), R \rightarrow R$, this output can be represented as follows:

$$G(\mathbf{x}; \boldsymbol{a}_i, b_i) = g(b_i + \boldsymbol{a}_i \cdot \mathbf{x}), b_i \in R \tag{2}$$

where $\boldsymbol{a}_i$ denotes the weight vector connecting the input layer to the ith hidden node, b_i denotes the bias of the ith hidden node, and $\boldsymbol{a}_i \cdot \mathbf{x}$ denotes the inner product of the vector $\mathbf{x}$ in R^n and $\boldsymbol{a}_i$. For RBF hidden nodes using a triangular activation function or the Gaussian function, $g(x) : G(\mathbf{x}; \boldsymbol{a}_i, b_i), R \rightarrow R$, the output can be represented as follows:

$$G(\mathbf{x}; \boldsymbol{a}_i, b_i) = g(\|\mathbf{x} - \boldsymbol{a}_i\| b_i), b_i \in R^+ \tag{3}$$

where $\boldsymbol{a}_i$ and b_i denote the impact factor and the center of the ith RBF node, respectively, and R^+ denotes all positive real values.

3.2 Extreme learning machines

An ELM is identical to a neural network with random weights, meaning that its learning process does not involve iterative tuning (Huang et al. 2006, 2012). The input weights and biases of hidden nodes are generated randomly, without

any correspondence to the data distribution, which results in a fast learning speed. This fast learning ability enables ELMs to be used in online graph classification (Han et al. 2015). Consider N arbitrary distinct samples $(\mathbf{x}_i, \mathbf{t}_i) \in R^n \times R^m, i = 1, 2, \ldots, N$, where $\mathbf{x}_i$ is an $n \times 1$ input vector, $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\mathrm{T}}$, and $\mathbf{t}_i$ is an $m \times 1$ "one-hot" vector, $\mathbf{t}_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^{\mathrm{T}}$. The output function of this network with L hidden nodes can be summarized as follows:

$$f_L(\mathbf{x}_j) = \sum_{i=1}^{L} \boldsymbol{\beta}_i G(\boldsymbol{a}_i, b_i, \mathbf{x}_j),\ j = 1, \ldots, N \tag{4}$$

where $\boldsymbol{a}_i = [a_{i1}, a_{i2}, \ldots, a_{iN}]$ is the weight vector connecting all input nodes to the ith hidden node and b_i is the bias of ith hidden node. These two parameters can be generated randomly at the very beginning of the model learning process (Huang et al. 2006). L is the number of hidden nodes, which is always determined by considering all possible values. Fortunately, Li et al. (2016) proposed a self-adaptive ELM that can self-adaptively determine the appropriate number of hidden nodes. $\boldsymbol{\beta}_i$ is the weight connecting the ith hidden node to the output layer. $G(\boldsymbol{a}_i, b_i, \mathbf{x}_j)$ is the output of the ith hidden node with respect to the input $\mathbf{x}_j$. $G(\boldsymbol{a}_i, b_i, \mathbf{x}_j)$ is a nonlinear piecewise continuous function (e.g., sigmoid or threshold function).

The goal of an ELM is to minimize the error between the real output and the expected output:

$$\begin{aligned} &\arg\min_{f} \| f_L(\mathbf{x}_j) - t_j \|^2 \\ &= \arg\min_{f} \left\| \sum_{i=1}^{L} \boldsymbol{\beta}_i G(\boldsymbol{a}_i, b_i, \mathbf{x}_j) - t_j \right\|^2,\ j = 1, \ldots, N \end{aligned} \tag{5}$$

Equation 5 can be summarized as

$$\arg\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|^2 \tag{6}$$

where

$$\begin{aligned} &\mathbf{H}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_L, b_1, \ldots, b_L, \mathbf{x}_1, \ldots, \mathbf{x}_L) \\ &= \begin{bmatrix} G(\boldsymbol{a}_1, b_1, \mathbf{x}_1) & \cdots & G(\boldsymbol{a}_L, b_L, \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ G(\boldsymbol{a}_1, b_1, \mathbf{x}_N) & \cdots & G(\boldsymbol{a}_L, b_L, \mathbf{x}_N) \end{bmatrix} \end{aligned} \tag{7}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\beta}_L^{\mathrm{T}} \end{bmatrix}_{L \times m} \quad and \quad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{t}_N^{\mathrm{T}} \end{bmatrix}_{N \times m} \tag{8}$$

The GLS solution for the ELM is

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T} \tag{9}$$

where $\mathbf{H}^{\dagger}$ is the Moore–Penrose generalized inverse of the matrix $\mathbf{H}$ (Serre 2002; Rao and Mitra 1972). It can be proven that the minimum error and the least norm are guaranteed to occur simultaneously (Huang et al. 2006). The highest output of $\mathbf{H}\boldsymbol{\beta}$ is chosen as the prediction result; this has been demonstrated to be reasonable in Wong et al. (2016).

Since ELM has a great advantage over the gradient-based methods in learning rate and generalizing learning ability, in recent years, different versions of improved ELM have been proposed. They have fully covered the problems of supervised learning (Zhang and Zhang 2017; Zhang and Deng 2017), semi-supervised learning (Liu et al. 2017), unsupervised learning (Ding et al. 2017) and even imbalance learning problems (Mao et al. 2017; Zhai et al. 2017). These articles have used a lot of ways to extend ELM to obtain higher prediction accuracy, better numerical stability and promising speed-up.

In addition to the above methods concerned only with the single domain problem, a series of extension methods based on ELM also have the cross-domain learning capability. For example, domain adaptation extreme learning machine (DAELM) can learn a robust classifier by leveraging a limited number of labeled data from target domain for drift compensation (Zhang and Zhang 2015). Cross-domain extreme learning machine (CdELM) can learn a common (shared) subspace across domains (Liu et al. 2017). And some other domain adaption ELM algorithms incorporated with manifold regularization (Zhang and Zhang 2016; Zhang et al. 2017) were proposed for visual categorization and image classification problem.

3.3 Semi-supervised manifold regularization

Manifold regularization refers to a geometric framework used for learning from labeled and unlabeled samples (Chai and Yang 2007). Methods of this kind depend on the following condition regarding the relation between the marginal and the conditional: if two points are close in the intrinsic geometry, then their conditional distributions are also similar. In other words, the conditional probability distribution varies smoothly along the geodesics in the intrinsic geometry.

For our case, this condition means that if two signal vectors $\boldsymbol{s}_{l_1}$ and $\boldsymbol{s}_{l_2}$ obtained from locations l_1 and l_2 are similar, as measured by the Euclidean norm $\|\boldsymbol{s}_{l_1} - \boldsymbol{s}_{l_2}\|$, then the conditional probabilities $P(l_1|\boldsymbol{s}_{l_1})$ and $P(l_2|\boldsymbol{s}_{l_2})$ should also be similar. More specifically, Pan et al. (2007) identified the following four characteristics of the relationship between the distributions of signal strength and physical location, as illustrated in Table 1.

Table 1 Signal strength (unit: dBm) (all values are rounded for illustration) (Pan et al. 2007)

	AP_1	AP_2	AP_3	AP_4	AP_5
t_A	− 40		− 60	− 40	− 70
t_B	− 50	− 60		− 80	
t_C		− 40	− 70		
t_D	− 80		− 40	− 70	
$t_{A'}$	− 40		− 70	− 40	− 60
t_E	− 40		− 70	− 40	− 80
t_F	− 80			− 80	− 50

- When a mobile device passes by the same location at two different times, e.g., t_A and $t_{A'}$, the pairwise signal strengths from most APs should be similar.
- If the received pairwise signal strengths from two APs are always similar, e.g., AP_1 and AP_4, then the locations of these two APs must be spatially close to each other.
- A mobile device and an AP may be spatially close if the signal is strong; e.g., if we place a mobile device immediately next to an AP, the signal can be as strong as -40 dBm.
- The positions of a mobile device at two consecutive times, e.g., $t_{A'}$ and t_E, should be spatially close to each other if the time interval between them is small, under the assumption that a user may not move too quickly or too irregularly.

Given all these assumptions, the optimal solution, according to Yang et al. (2015), will be $\boldsymbol{f}^* = \arg\min_{\boldsymbol{f}} \sum_{i=1}^{l} \|f_i - y_i\|^2 + \gamma \boldsymbol{f}^{\mathrm{T}} \boldsymbol{L} \boldsymbol{f}$, where the first term measures the fitting error with respect to the labeled data, the second term refers to the manifold, and $\boldsymbol{L}$ is the graph Laplacian of the high-dimensional signal space. Using this semi-supervised learning model, with only a small amount of labeled data, it's easy to obtain the coordinates of a smooth trajectory. Regarding the optimal solution, it is physically equivalent to a smooth trajectory that passes through as many labeled points as possible.

For all collected signal vectors $\boldsymbol{s}_t (t = 1, 2, \ldots, m)$, it first constructs the neighborhood graph and its graph Laplacian matrix $\boldsymbol{L}$ (Chung 1997). Then, the goal is to minimize the following objective function:

$$\mathbf{P}^* = \underset{\mathbf{P} \in R^{m \times 2}}{\arg\min} (\mathbf{P} - \mathbf{Y})^{\mathrm{T}} \boldsymbol{J} (\mathbf{P} - \mathbf{Y}) + \gamma \mathbf{P}^{\mathrm{T}} \boldsymbol{L} \mathbf{P} \tag{10}$$

Here, $\mathbf{P}$ is an $m \times 2$ matrix representing the positions of the mobile device that are to be determined; $\boldsymbol{J} = diag(\delta_1, \delta_2, \ldots, \delta_m)$ is an indication matrix, where $\delta_t = 1$ if the coordinates of the mobile device at time t are given and $\delta_t = 0$ otherwise; $\mathbf{Y} = [\mathbf{y}_1^{\mathrm{T}}, \mathbf{y}_2^{\mathrm{T}}, \ldots, \mathbf{y}_m^{\mathrm{T}}]^{\mathrm{T}}$ is an $m \times 2$ matrix representing the calibration data, where $\mathbf{y}_t$ represents the given coordinates of the mobile device at time t if $\delta_t = 1$ and otherwise can take any set of values, e.g., $\mathbf{y}_t = [0\ 0]$; and γ controls the smoothness of the coordinates along the manifold.

By setting the derivative of Eq. 10 to zero, the optimal solution is obtained as follows Ham et al. (2005):

$$\mathbf{P}^* = (\boldsymbol{J} + \gamma \boldsymbol{L})^{-1} \boldsymbol{J} \mathbf{Y} \tag{11}$$

4 FSELM: fusion semi-supervised extreme learning machine

In this paper, we propose an indoor localization method with high accuracy and a low human calibration cost. To this end, we extend the ELM approach to FSELM in two respects: (1) we incorporate semi-supervised manifold regularization to reduce the dependence on human calibration, and (2) we exploit multiple signals to improve the location accuracy. Here, we consider Wi-Fi and BLE signals because various buildings, such as shopping malls, hospitals, schools and airports, are commonly equipped with the necessary infrastructure.

Unlike in previous single-signal-based semi-supervised manifold methods (Liu et al. 2011; Pan et al. 2006, 2007; Yang et al. 2015; Zhang and Zhi 2010; Scardapane et al. 2016), we combine the Wi-Fi and BLE signals into a single model. To the best of our knowledge, Wi-Fi and BLE signals have different propagation characteristics and effective distances. Figure 1 shows the neighbor relations of the signal vectors for Wi-Fi and BLE. Every vertex in Fig. 1 denotes a position at which one Wi-Fi or BLE RSS vector is obtained, and a link between two vertexes indicates that the corresponding RSS vectors are similar to each other. Figure 1a shows the neighbor relations for Wi-Fi, and Fig. 1b shows the neighbor relations for BLE. Because Wi-Fi has a larger effective distance than BLE, it also has more complex neighbor relations; e.g., there is a link between vertex "3" and vertex "8" for Wi-Fi signals, but no such link exists for BLE. Therefore, when we wish to consider both Wi-Fi and BLE in a semi-supervised learning model, we should build the manifold regularization separately for each of them.

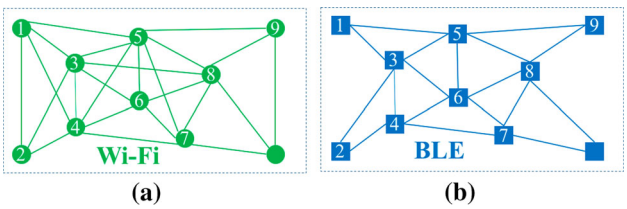

Fig. 1 Similarity relations of signal samples: **a** Wi-Fi and **b** BLE (a line indicates that two samples are neighbors along the manifold)

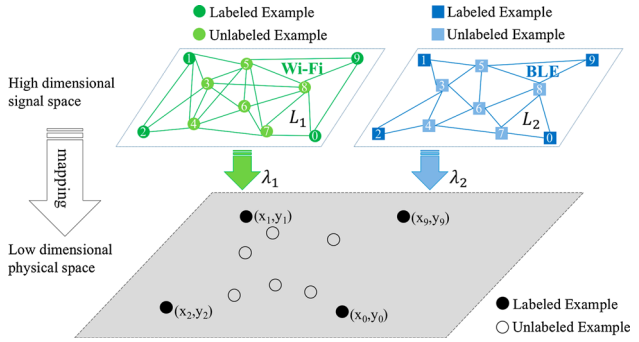

Fig. 2 FSELM model illustration (L_1 and L_2 are the graph Laplacians of the Wi-Fi and BLE signals, and λ_1 and λ_2 are the weight coefficients of the two manifold constraints)

In accordance with the structural risk minimization principle (Vapnik 2013), FSELM uses graph Laplacian regularization to find the structural relationships of both the labeled and unlabeled samples in the high-dimensional signal space. For the construction of a semi-labeled graph G based on l labeled samples and u unlabeled samples, each collected signal vector $\boldsymbol{s}_j = [s_{j1}, s_{j2}, \ldots, s_{jN}] \in R^N$ is represented by a vertex j, and if the vertex j is one of the neighbors of i, we represent this by drawing an edge with a weight of w_{ij} connecting them. According to Belkin et al. (2006), the graph Laplacian $\boldsymbol{L}$ can be expressed as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W}$. Here, $\boldsymbol{W} = [w_{ij}]_{(l+u)\times(l+u)}$ is the weight matrix, where $w_{ij} = exp(-\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2/2\sigma^2)$ if $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ are neighbors along the manifold and $w_{ij} = 0$ otherwise, and $\boldsymbol{D}$ is a diagonal matrix given by $\boldsymbol{D}_{ii} = \sum_{j=1}^{l+u} \boldsymbol{W}_{ij}$.

As illustrated in Fig. 2, to consider the empirical risk while controlling the complexity, FSELM minimizes the fitting error plus two separate smoothness penalties for Wi-Fi and BLE as follows:

$$\arg\min_{\boldsymbol{f}} \left\{ \frac{1}{2}\|\boldsymbol{f} - \mathbf{T}\|^2 + \lambda_1 \boldsymbol{f}^{\mathrm{T}} \boldsymbol{L}_1 \boldsymbol{f} + \lambda_2 \boldsymbol{f}^{\mathrm{T}} \boldsymbol{L}_2 \boldsymbol{f} \right\} \tag{12}$$

The first term represents the empirical error with respect to the labeled training samples, and the second and third terms represent the manifold constraints for Wi-Fi and BLE based on the graph Laplacians $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$. By adjusting the two coefficients λ_1 and λ_2, we can control the relative influences of the Wi-Fi and BLE signals on the model. By substituting the underlying ELM solution $\boldsymbol{f} = \mathbf{H}\boldsymbol{\beta}$ into Eq. 12, we obtain

$$\begin{aligned} \arg\min_{\boldsymbol{\beta}} l(\boldsymbol{\beta}) = \arg\min_{\boldsymbol{\beta}} \Big\{ & \frac{1}{2}\|\boldsymbol{J}\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|^2 \\ & + \lambda_1 (\mathbf{H}\boldsymbol{\beta})^{\mathrm{T}} L_1 \mathbf{H}\boldsymbol{\beta} + \lambda_2 (\mathbf{H}\boldsymbol{\beta})^{\mathrm{T}} L_2 \mathbf{H}\boldsymbol{\beta} \Big\} \end{aligned} \tag{13}$$

where $\mathbf{H}$ is still the output matrix, with dimensions of $(l+u) \times \widetilde{N}$ because all of the labeled and unlabeled data are considered here, and $\boldsymbol{J} = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_{l+u})$ is an indication matrix, where $\delta_i = 1$ if the ith sample is labeled and $\delta_i = 0$ otherwise. The length of the target vector $\mathbf{T}$ in SELM is extended to $l+u$, where l elements of $\mathbf{T}$ correspond to the real data and the remaining u elements are equal to 0.

For the optimization problem expressed in Eq. 13, we arrive at the following convex differentiable objective function:

$$\begin{aligned} \frac{\partial l}{\partial \boldsymbol{\beta}} = 0 \Rightarrow & (\boldsymbol{J}\mathbf{H}\boldsymbol{\beta} - \mathbf{T})^{\mathrm{T}} \boldsymbol{J}\mathbf{H} + \lambda_1 (\mathbf{H}\boldsymbol{\beta})^{\mathrm{T}} \boldsymbol{L}_1 \mathbf{H} \\ & + \lambda_2 (\mathbf{H}\boldsymbol{\beta})^{\mathrm{T}} \boldsymbol{L}_2 \mathbf{H} = 0 \end{aligned} \tag{14}$$

Finally, $\boldsymbol{\beta}$ can be derived as follows:

$$\boldsymbol{\beta} = \left(\left(\boldsymbol{J} + \lambda_1 \boldsymbol{L}_1^{\mathrm{T}} + \lambda_2 \boldsymbol{L}_2^{\mathrm{T}} \right) \mathbf{H} \right)^{\dagger} \boldsymbol{J}\mathbf{T} \tag{15}$$

It is obvious that if either λ_1 or λ_2 in Eqs. 14 and 15 is set to zero, the solution for $\boldsymbol{\beta}$ will be the same as in Liu et al. (2011). In this sense, FSELM can be regarded as an extension of SELM. Moreover, if both λ_1 and λ_2 are set to zero, the model will degenerate to an ELM.

To summarize, we consider a dataset that includes labeled samples $\{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in R^{n_1+n_2}, \mathbf{t}_i \in R^m, i = 1, 2, \ldots, l\}$ and unlabeled samples $\{(\mathbf{x}_i) | \mathbf{x}_i \in R^{n_1+n_2}, i = 1, 2, \ldots, u\}$, where $\mathbf{x}_i = [\mathbf{x}_i^{'}\ \mathbf{x}_i^{''}]$ consists of n_1 BLE RSS measurements and n_2 Wi-Fi RSS measurements. Some of the parameters are specified, such as the number of hidden nodes $\widetilde{N}$ and the constraint weights λ_1 and λ_2. In addition, the activation function $g(\mathbf{x})$ and the node type ("RBF") are selected. Then, the FSELM algorithm can be summarized into the following five steps:

1. Randomly assign input weights $\boldsymbol{a}_i$ and biases b_i, $i = 1, 2, \ldots, \widetilde{N}$, where $\widetilde{N}$ is the number of hidden nodes.
2. Calculate the hidden-layer output matrix $\mathbf{H}$ using all labeled and unlabeled samples; here, $\mathbf{H}$ is an $(l+u) \times \widetilde{N}$ matrix.
3. Separately calculate the graph Laplacians $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ for BLE and Wi-Fi, respectively. $\boldsymbol{L}_1 = \boldsymbol{D}^{'} - \boldsymbol{W}^{'}$, $\boldsymbol{D}_{ii}^{'} = \sum_{j=1}^{l+u} \boldsymbol{W}_{ij}^{'}$, $\boldsymbol{W}_{ij}^{'} = \sum_{j=1}^{l+u} exp(-|\mathbf{x}_i^{'} - \mathbf{x}_j^{'}|^2/2\delta^2)$, $\boldsymbol{L}_2 = \boldsymbol{D}^{''} - \boldsymbol{W}^{''}$, $\boldsymbol{D}_{ii}^{''} = \sum_{j=1}^{l+u} \boldsymbol{W}_{ij}^{''}$, and $\boldsymbol{W}_{ij}^{''} = \sum_{j=1}^{l+u} exp(-|\mathbf{x}_i^{''} - \mathbf{x}_j^{''}|^2/2\delta^2)$, where $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are both $(l+u) \times (l+u)$ matrixes.
4. Calculate $\boldsymbol{\beta} = \left(\left(\boldsymbol{J} + \lambda_1 \boldsymbol{L}_1^{\mathrm{T}} + \lambda_2 \boldsymbol{L}_2^{\mathrm{T}} \right) \mathbf{H} \right)^{\dagger} \boldsymbol{J}\mathbf{T}$.
5. Use the model $\boldsymbol{f} = \mathbf{H}\boldsymbol{\beta}$ to generate the prediction.

5 Experimental evaluation

To evaluate the performance of our proposed FSELM, we conducted experiments in two realistic indoor scenes, which are well instrumented with Wi-Fi and BLE APs. In this section, we first describe the experimental conditions and

Table 2 Dataset description

Dataset	Size (m^2)	#samples	#BLE	#Wi-Fi
Mall	31.8×14.95	9492	22	35
Office	12.5×7.5	3916	20	30

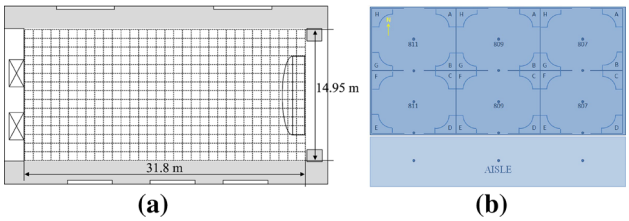

Fig. 3 The floorplan of the experimental site: **a** is a lobby of a shopping mall; **b** is a office area

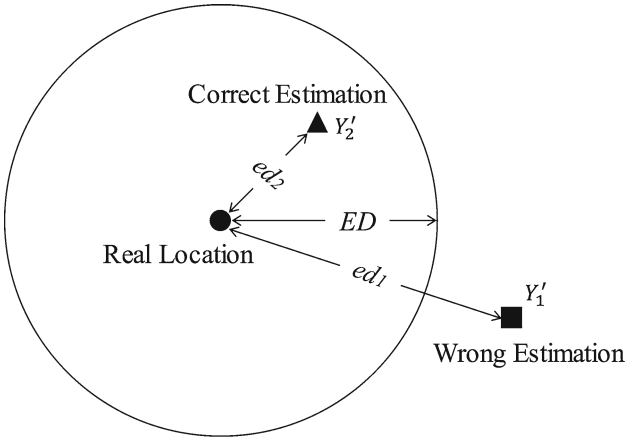

Fig. 4 The evaluation of a predicted result based on ed (error distance): Y_1' is regarded as a wrong location result as $ed_1 > ED$; but Y_2' is correct as $ed_2 < ED$

data acquisition. Afterward, comparisons with many previous methods are presented.

5.1 Experimental conditions introduction

5.1.1 Data acquisition

In our experiments, all algorithms were run on a desktop computer with (1) the Windows 7 Ultimate x64 operating system, (2) an Intel(R) Core(TM) i5-4200M @2.50 GHz CPU, (3) 4 GB of RAM and (4) MATLAB R2013a as the simulation software.

We evaluated the proposed method in two realistic indoor environment using real data collected by smartphones. As shown in Table 2, one dataset was collected in the lobby of Jinyuan Yansha Mall in Beijing, China, which has dimensions of $31.8 \times 14.95\,m^2$. The floorplan is shown in Fig. 3a. We divided the lobby into 480 square grid cells, each with an area of $1\,m^2$, and each grid cell was designated as an acquisition point. We collected 20 samples in each grid cell; thus, 9600 samples were obtained. However, we discarded 108 samples for being in the wrong format. Thus, the final dataset contained 9492 samples. Around the locations of all acquisition points, a total of 22 BLE APs and 35 Wi-Fi APs are present. The other dataset was collected in a office area, which has the dimensions of $12.5 \times 7.5\,m^2$. The floorplan is shown in Fig. 3b. We have set up 104 acquisition points evenly and in total collected 3916 samples. Each sample includes two-dimensional coordinates (x, y) and 50 signal strength features (20 BLE and 30 Wi-Fi). Any missing RSS values were set to -95 dBm, which is the minimum strength of the signals received in the test environment according to Chen et al. (2006).

5.1.2 Location accuracy measurement

Unlike in other conventional classification and regression problems, in location estimation, the location accuracy will vary under different error distance conditions. Given a well-trained location model f, for a testing sample $\{X_t, Y_t\}$, where X_t is the feature vector and Y_t is the true position, we can obtain a predicted position $Y' = f(X_t)$. As shown in Fig. 4, the prediction error distance between Y' and Y_t can be calculated using the "Euler Distance" formula $ed = \|Y' - Y_t\|_2$. If ed is less than ED, the tolerance threshold for the error distance, then the result is regarded as a correct one; otherwise, it will be regarded as incorrect. In the following illustration, all location accuracies will be discussed with respect to a certain error distance ED.

In the following, we will first fully verify the performance of the algorithm on the "Mall" dataset and then carry out similar validation on the "Office" dataset. The verification tests include: (1) the semi-supervised learning ability in the case that only part of the training data is labeled; (2) the generalization ability on the testing dataset; and (3) the time consumption for model training and testing.

5.2 Performance on "Mall" dataset

5.2.1 Model parameter selection

For the ELM, SELM and FSELM algorithms, the number of hidden nodes and the activation function need to be specified. Referring to the experimental conditions reported in Huang et al. (2012), we set the number of hidden nodes to 1000. In addition, the nonlinear sigmoid function ('sig') was selected as the activation function. The manifold constraint coefficients, λ in SELM and (λ_1, λ_2) in FSELM, are empirical parameters that need to be determined through experiments.

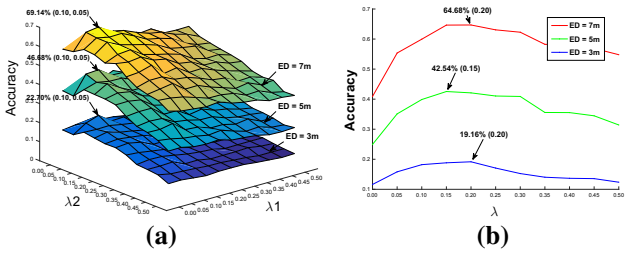

Fig. 5 Model parameter selection (selection of the best parameters λ_1 and λ_2 for FSELM and the best λ for SELM)

Table 3 Parameters of the semi-supervised manifold learning methods for "Mall" dataset

Method	λ	#nodes	$g(\cdot)$
LeMan	$\lambda = 1$	–	–
FSELM	$\lambda_{ble} = 0.1, \lambda_{\text{wifi}} = 0.05$	1000	'sig'
SELM	$\lambda = 0.2$	1000	'sig'
FS-RVFL	$\lambda_{\text{ble}} = 0.4, \lambda_{\text{wifi}} = 0.4$	1000	'sig'
S-RVFL	$\lambda = 0.9$	1000	'sig'

From the "Mall" dataset, we used 10 samples from each grid cell as the training dataset, for a total of 4800 samples. The remaining 4692 samples were used as the testing data. Of the 4800 training samples, we randomly labeled 500 samples with their real position coordinates, in the form $\{X_l, Y_l\}$, and left the rest as unlabeled data, in the form $\{X_u\}$.

For FSELM, we obtained the optimal pair of (λ_1, λ_2) values through grid search. For both λ_1 and λ_2, 11 different values were tested: $\{0, 0.05, 0.1, \ldots, 0.5\}$. For SELM, we tested the location accuracy with λ values selected from $\{0, 0.05, 0.1, \ldots, 0.5\}$. The range of 0 to 0.5 was selected because a weight should be a positive number and experimental results show that the location accuracy sharply declines once λ is larger than 0.5. The experimental results are shown in Fig. 5.

Figure 5a presents the results for FSELM with different values of λ_1 and λ_2. We calculated the location accuracies for ED values of 3 m, 5 m and 7 m. Since both λ_1 and λ_2 had 11 possible values between 0 and 0.5, we obtained 121 (11×11) location results for each error distance condition. From Fig. 5a, we can see that as λ_1 and λ_2 increase, the location accuracy will initially increase and later decline, and the global optimal location result is obtained when (λ_1, λ_2) is (0.10, 0.05). Similarly, Fig. 5b indicates that when ED is 3 and 7 m, SELM will achieve the optimal accuracy when λ is 0.2, whereas for $ED = 5$ m, the optimal location result is obtained when λ is equal to 0.15 or 0.2. Therefore, in the following experiments, we set λ to 0.2 for the SELM method. These results suggest that an appropriate coefficient can make full use of the high-dimensional signal-space manifold information to improve the location accuracy. However, excessively large coefficients will result in overfitting of the signal information contained in unlabeled data. Therefore, in practical applications, we need to determine the appropriate value(s) of the manifold constraint coefficient(s) to ensure a good predictive effect.

5.2.2 Semi-supervised manifold learning ability

As a semi-supervised learning method, our proposed FSELM relies on manifold regularization to build a location model from both labeled and unlabeled data. We compare FSELM with several previous semi-supervised learning methods: LeMan (Pan et al. 2006), SELM (Liu et al. 2011), FS-RVFL and S-RVFL. LeMan is a manifold regularization approach to calibration reduction for sensor network-based tracking. It solves the problem of tracking the trajectory of a person walking through a predefined area equipped with a sensor network when only some of the coordinates along the trajectory are known.

The RVFL method relies on the learning and generalization characteristics of random vector functional link net. The difference between the ELM and RVFL approaches is that RVFL considers not only links from the enhanced pattern to the outputs but also direct links from the inputs to the outputs. Based on the RVFL approach, we constructed two variants: semi-supervised RVFL (S-RVFL) and fusion semi-supervised RVFL (FS-RVFL). The parameters of the methods mentioned above are shown in Table 3.

For LeMan, SELM and S-RVFL, there is only one coefficient, λ, because the Wi-Fi and BLE signal strength values are combined into one vector. For FSELM and FS-RVFL, there are two coefficients, λ_{ble} and λ_{wifi}, because the Wi-Fi and BLE graph Laplacians are calculated separately. The optimal coefficient values listed in Table 3 were selected using the method described in Sect. 5.2.1. In our experiments, we set the number of hidden nodes to 1000 and selected the nonlinear sigmoid function ('sig') as the activation function.

We compared the performance of the various methods on the 4800 preselected training samples while varying the number of labeled samples N_l from 50 to 4500. For LeMan, the manifold learning ability can be assessed based on the regression accuracy, whereas the other four approaches should be evaluated in terms of the training accuracy. We show the results for 9 selected values of N_l in Fig. 6. As N_l increases, the accuracy increases for all five methods. For example, the accuracy when $N_l = 4000$ is almost 40% higher than it is when $N_l = 100$ under the condition of 5 m error distance. Moreover, the FSELM and FS-RVFL methods perform similarly to each other and achieve the highest accuracy, indicating that for multi-signal-based localization, it is better to consider each signal as an independent Laplacian constraint. Meanwhile, FSELM shows somewhat better performance

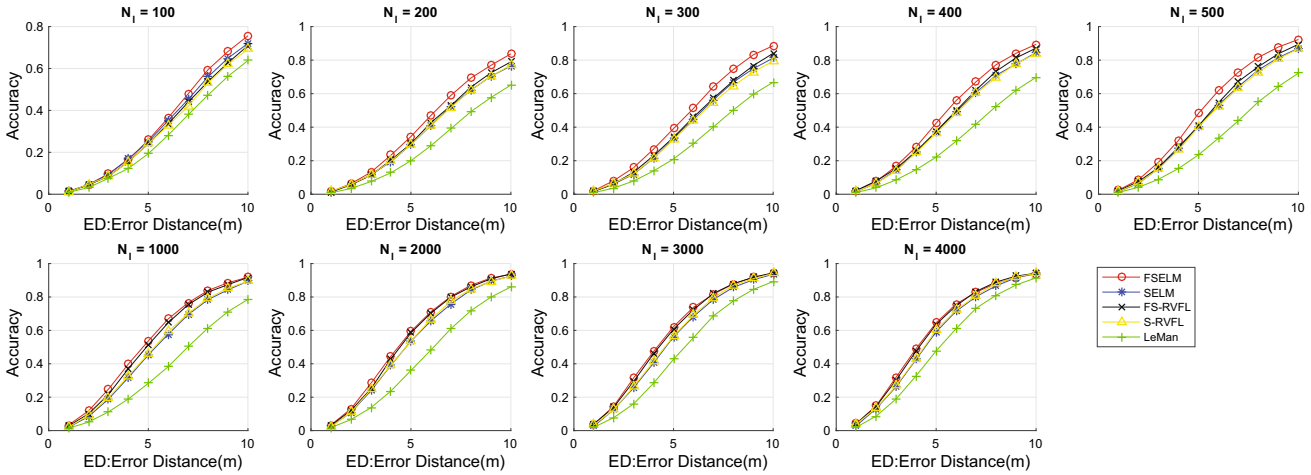

Fig. 6 Semi-supervised manifold learning performance on "Mall" dataset (the training accuracies comparison)

than FS-RVFL for two reasons. First, FSELM considers only a nonlinear mapping from the inputs to the outputs, whereas RVFL also considers direct links from the inputs to the outputs. For the problem of interest here, direct links are less effective than a nonlinear mapping for mining the nonlinear relationship between the signal inputs and the physical coordinate outputs. Second, the performance strongly depends on the values of the manifold constraint coefficients, and the λ values that we selected are not suitable for all different values of N_1. Notably, among the five methods, LeMan is less effective than the others because LeMan considers only manifold regularization, whereas the other four methods also create a mapping function between the signals and the physical locations by constructing a network with a single hidden layer, which helps to improve their accuracy.

When analyzing these algorithms at the theoretical level, the mapping accuracy from the signal space to the physical coordinate constructed by LeMan is strongly dependent on the manifold constraints. In the case of sparse calibration, the local errors of manifold constraints will cause great errors in location accuracy. But for the other four methods, besides the manifold constraints, they also have the nonlinear mapping relation constructed by ELM or RVFL model. Therefore, the robustness of the algorithm is more obvious when the calibration data is sparse. Moreover, constructing the manifold constraints for Wi-Fi and BLE respectively, in fact, is a strategy to integrate manifold constraints of varying degrees, so as to further ensure the accurate neighbor relationship. Therefore, the FSELM model has formed a complementary relationship from several aspects, thus achieving more ideal results.

5.2.3 Location accuracy comparison

For the location accuracy comparison, we compared the testing accuracies of FSELM with those of the SELM, FS-RVFL, S-RVFL, ELM, BP and SVM methods for different proportions of calibration data to evaluate the performance of the proposed method.

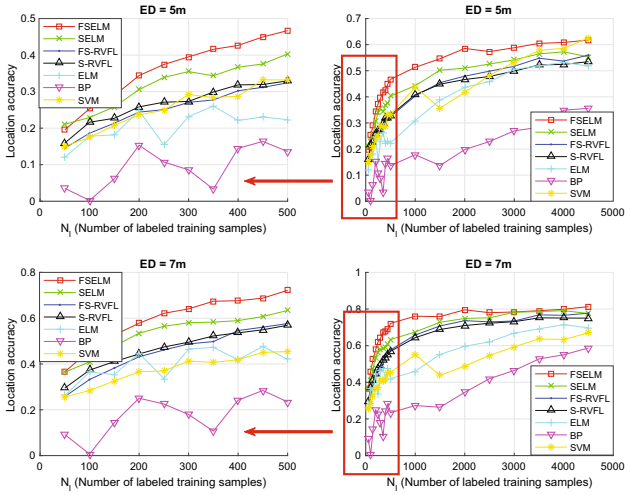

Fig. 7 Location accuracies on "Mall" dataset under conditions of $ED = 5$ m and $ED = 7$ m as the number of labeled training samples increases from 100 to 4500 (the testing accuracies comparison)

We conducted localization experiments corresponding to $ED = 5$ m and $ED = 7$ m. We again used 4800 samples as training data and evaluated the location accuracy on the remaining 4692 testing data. The four semi-supervised methods (SELM, FSELM, S-RVFL and FS-RVFL) can use both labeled and unlabeled training data, whereas ELM, BP and SVM can only use labeled samples for model training. Figure 7 shows the regression accuracies for different numbers of labeled samples, from 50 to 4500.

From Fig. 7, we can conclude that as the amount of labeled calibration data increases, all seven models can achieve increasingly high accuracy. However, FSELM performs better than the other six models with the same amount of labeled

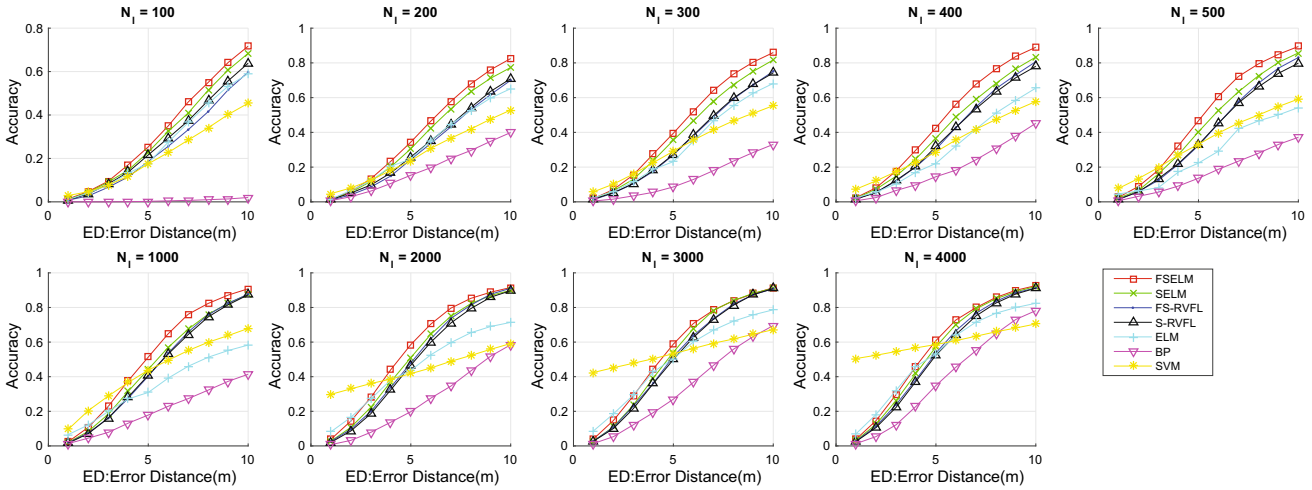

Fig. 8 Location accuracies of the seven methods under varying ED conditions with different numbers of labeled training samples $N_\mathrm{l} = \{100, 200, 300, 400, 500, 1000, 2000, 3000, 4000\}$ on "Mall" dataset. The total number of labeled and unlabeled training samples is $N_\mathrm{l}+N_\mathrm{u} = 4800$. The number of testing samples is 4692

data, thereby demonstrating its good performance in location estimation. FSELM, SELM, FS-RVFL and S-RVFL all achieve better location accuracies than the ELM, BP and SVM methods because in addition to using the labeled calibration data, these semi-supervised learning methods can also extract useful information from the large amount of available unlabeled data to improve their location accuracy. Especially with fewer than 500 labeled training samples, the former methods achieve higher and more stable location accuracies. By contrast, the location accuracies of the ELM and BP methods are low and unstable because there are too few labeled samples to train a model with a good generalization ability. However, the SVM method still performs well enough, as indicated by its good generalization ability. Figure 7 also reveals that in comparison with SELM, FSELM can achieve higher location accuracy even with the same numbers of labeled and unlabeled samples. Therefore, we can conclude that with regard to the manifold constraints, it is much better to consider the Wi-Fi and BLE signals separately than jointly.

From Fig. 7, we can also sum up the advantages of the FSELM method on reducing human calibration. For example, under the error distance of 5 meters, FSELM can achieve the location accuracy of 58% with 2000 labeled samples. But other methods need more than 4000 labeled samples to achieve the same accuracy. And for the situation of 7 meters error distance, it is more obvious that as the labeled samples increasing, FSELM is the first method to achieve 80% accuracy when N_l equals 2000, and the second method, SELM cannot achieve this accuracy until N_l is more than 4500. This advantage on reducing human calibration makes FSELM easier to be applied in realistic problems.

In addition, to further compare the location accuracies of the seven methods mentioned above, we selected 100, 200, 300, 400, 500, 1000, 2000, 3000 and 4000 labeled training samples. Figure 8 shows the relationship between the location accuracy and the error distance for each of these calibration data conditions.

From Fig. 8, we can draw similar conclusions: (1) The semi-supervised learning methods (FSELM, SELM, FS-RVFL and S-RVFL) can use the unlabeled samples to further improve their location accuracies compared with the supervised models (ELM, BP and SVM) given the same number of labeled samples. Because supervised learning models are often less fitting on the sparse data, which will directly lead the low and unstable accuracy. (2) For manifold learning, it is better to build two graph Laplacian constraints for Wi-Fi and BLE signals separately than to build a single graph Laplacian constraint for the Wi-Fi and BLE signal strength values jointly. Because this different granularity mapping relationship can often form complementarity, thus improving its robustness. (3) Moreover, as the number of labeled samples increases, the accuracy of the ELM method will increase to a level similar to that for FSELM and SELM; however, BP still cannot reach very high accuracy. This suggests that the ELM-based models possess a better learning ability than that of BP for this location problem. (4) The ELM and SVM methods have a good generalization ability, thereby ensuring the stability of their accuracy performance.

5.2.4 Time consumption

As mentioned previously, the time consumption for training and testing is also a very important concern for model deployment in practical applications. For a constant training set size

of 4800, we evaluated the time consumption while varying the number of labeled samples N_l from 500 to 4500. Figure 9 shows the results for all seven models, in which, the color bars of ELM, SVM and BP show the total time consumption for model training. But for FSELM, FS-RVFL, SELM and S-RVFL, the colors parts only indicate the time consumption for model optimization. And the dark parts indicate the time for calculating the manifold regularization. The reason why the dark parts of FSELM and FS-RVFL are almost two times of SELM and S-RVFL is that there are two manifold constraints in FSELM and FS-RVFL, but there is only one manifold constraint in SELM and S-RVFL.

The batch learning approaches (ELM, BP and SVM) show a tendency for their training time consumption to increase with increasing N_l because they use only labeled data for model training, and more labeled samples will require more time to process. By contrast, the four semi-supervised approaches (FSELM, SELM, FS-RVFL and S-RVFL) maintain a nearly constant time consumption because their training time depends only on the total number of training samples. For this reason, when N_l is much smaller than the total number of training samples, the training time consumption of the semi-supervised learning approaches is much higher than that of the batch learning approaches. In addition, the time consumption of manifold constraint is mainly due to the operation of large matrix, which can be further reduced by matrix decomposition and parallel computing.

Meanwhile, Fig. 9 shows the testing times on the 4692 testing samples. Unlike the other six approaches, the SVM method requires an increasingly longer time for testing as N_l increases, even though the number of testing samples never changes. We believe that the main reason for this behavior might be the model complexity. As the number of training samples increases, the SVM method will construct an increasingly complex model to minimize the fitting error. Consequently, even for the testing task, this greater complexity will result in a higher time cost.

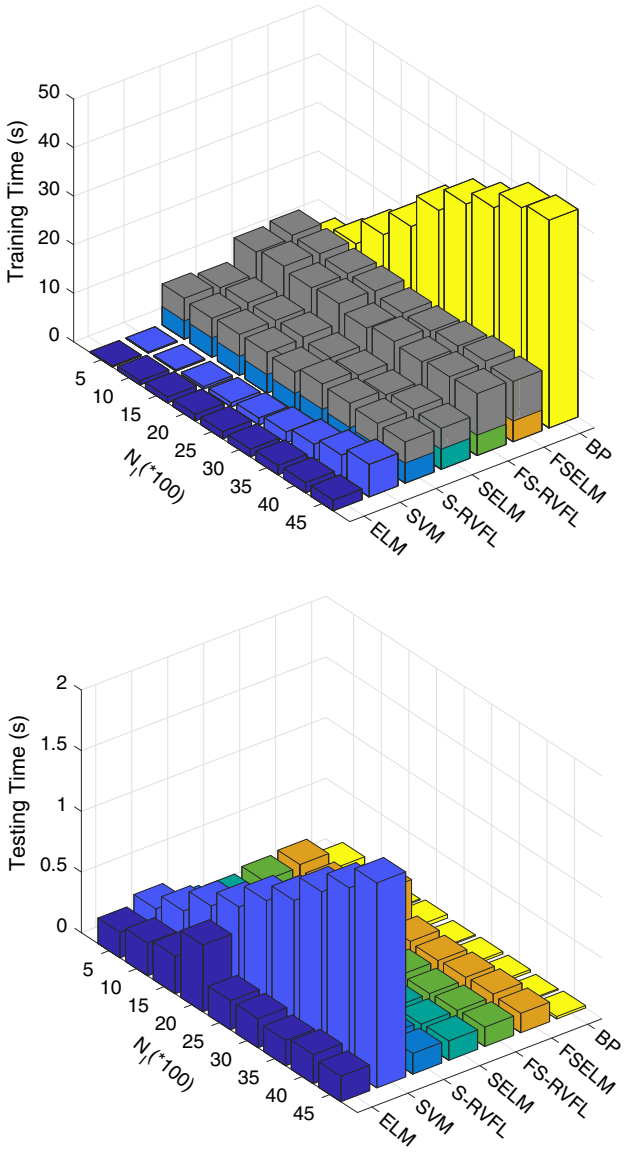

Fig. 9 Training and testing time consumption results for the seven methods on "Mall" dataset. There are 4800 training samples, of which 500–4500 are labeled, and the number of testing samples is 4692

5.3 Performance on "Office" dataset

5.3.1 Model parameter selection

Like the experiments for "Mall" dataset, we also optimize the parameters of all models for "Office" dataset. The number of hidden nodes is still set to be 1000, and the nonlinear sigmoid function ('sig') is still adopted as the activation function. The manifold constraint coefficients are determined through the cross-validation experiments on the whole dataset. We used the average value of location accuracy under the error distance of 1–5 m to make the global optimal parameters choices, shown in Table 4.

Table 4 Parameters of the semi-supervised manifold learning methods for "Office" dataset

Method	λ	#nodes	$g(\cdot)$
LeMan	$\lambda = 1$	–	–
FSELM	$\lambda_{ble} = 0.2, \lambda_{wifi} = 0.05$	1000	'sig'
SELM	$\lambda = 0.3$	1000	'sig'
FS-RVFL	$\lambda_{ble} = 0.35, \lambda_{wifi} = 0.4$	1000	'sig'
S-RVFL	$\lambda = 0.7$	1000	'sig'

5.3.2 Manifold learning accuracy and location accuracy

To evaluate the manifold learning accuracy and location accuracy, we randomly select 3000 samples as the training

Table 5 Manifold learning accuracy and location accuracy on "Office" dataset with $ED = 2$ and 3 m

N_l	100	300	500	1000	1500	2000	2500	3000
Training accuracy % (2m/3m)								
FSELM	**48.7/74.6**	**66.5/87.8**	**69.0/89.9**	**76.2/93.7**	**79.3/95.0**	**81.1/97.2**	**81.1/97.2**	**83.0/97.3**
FS-RVFL	24.6/50.9	55.3/79.3	63.3/88.0	73.4/92.8	77.1/*95.7*	76.5/94.5	80.4/95.6	82.6/96.8
SELM	35.6/60.7	51.6/75.9	60.1/84.1	70.2/90.7	75.8/94.6	79.2/95.0	77.1/95.4	76.8/94.7
S-RVFL	20.3/41.2	32.9/60.7	41.8/68.7	53.8/80.3	58.7/85.6	63.4/89.5	65.0/91.0	69.2/94.2
LeMan	17.1/35.4	25.5/47.5	29.9/56.5	43.3/73.1	52.5/82.3	59.0/87.0	64.0/90.0	67.7/91.7
Testing accuracy % (2m/3m)								
FSELM	**45.3/73.5**	**59.2/83.2**	**62.8/84.9**	**69.8/92.7**	**74.0/93.7**	**75.7/94.1**	**77.1/94.8**	**78.6/95.5**
FS-RVFL	24.2/46.9	49.7/79.5	57.9/83.3	69.3/89.0	72.1/91.7	73.5/92.5	72.5/93.0	74.9/92.6
SELM	35.8/59.0	49.5/74.5	54.8/80.8	65.2/87.8	68.0/90.6	71.2/91.8	71.0/89.6	75.7/92.9
S-RVFL	22.1/42.1	35.7/61.0	40.7/68.0	52.6/80.0	57.4/83.5	62.0/88.1	64.6/90.4	67.6/90.9
ELM	4.9/ 8.5	38.0/61.4	47.3/71.8	63.2/84.1	61.5/84.6	62.9/84.3	65.9/88.0	68.9/90.6
BP	29.8/50.3	38.7/68.6	43.0/71.5	47.9/78.6	48.0/80.5	52.2/81.4	52.8/82.3	55.7/84.7
SVM	28.5/44.2	28.5/44.4	37.6/53.6	45.7/66.9	50.8/67.3	57.4/74.7	59.7/77.1	61.6/79.3

Bold values indicate the best accuracy

Table 6 Training and testing time consumption for "Office" dataset

N_l	1000	2000	3000
Training time (s)			
FSELM	5.75+3.39	5.69+3.13	5.08+3.16
SELM	5.27+1.53	5.56+1.89	5.53+1.31
FS-RVFL	5.03+3.16	5.39+3.25	4.95+3.05
S-RVFL	4.92+1.81	5.55+1.67	4.84+1.72
ELM	1.59	3.02	3.80
BP	18.14	26.03	38.61
SVM	0.52	6.17	11.38
Testing time (s)			
FSELM	0.19	0.17	0.16
SELM	0.09	0.19	0.14
FS-RVFL	0.16	0.13	0.25
S-RVFL	0.23	0.22	0.16
ELM	0.16	0.23	0.20
BP	0.02	0.06	0.02
SVM	0.58	1.08	1.48

dataset while setting N_l of them as the labeled samples and keeping the rest N_u as unlabeled samples. While N_l increasing from 100 to 3000, the manifold learning accuracies of FSELM, FS-RVFL, SELM, S-RVFL, LeMan are shown in Table 5. We can find that for the most condition of N_l, FSELM can get the best accuracy, when there are more then 2500 labeled samples, both FSELM and FS-RVFL achieve superior accuracy. This is because these two methods fully take use of the fusion signals via separate manifold conditions of BLE and Wi-Fi. However, when N_l is less than 300, FSELM shows the absolute excellence in learning ability. We think the reason is that in the office environment, the nonlinear mapping between the signal feature and physical location is more significant, so the FSELM model is still very robust under very few calibration cases.

Besides the manifold learning ability, FSELM also has excellent performance in location prediction comparing with other six methods. Besides, FS-RVFL, SELM and S-RVFL can also reach relatively high accuracy, which indicates that manifold constraint is a powerful auxiliary item for model training besides fitting error. In addition, we note that with the calibration data increasing, ELM's accuracy rises very fast, but in comparison, BP and SVM are a little worse, it also shows ELM's strong learning ability.

5.3.3 Time consumption

To validate models' learning speed, we test the training time on a dataset with 3000 samples, in which the labeled samples N_l vary from 1000 to 3000, and test the testing time on another dataset with 3000 samples. From Table 6, we can find that the computation of each Laplacian matrix will cost about 1.5 s in SELM and S-RVFL. FSELM and FS-RVFL need about 3.3 s as they have two Laplacian manifold terms. And these four semi-supervised algorithms almost maintain the same time consumption for optimization objective solution, as it related to the total training samples. And three semi-supervised methods' training time is determined by the actual amount of labeled training samples. The time consumption of the testing process is very small, mainly because the prediction process of these models only contains the basic matrix operation.

6 Conclusion and future work

In this paper, we have proposed a fusion semi-supervised extreme learning machine model called FSELM for data fusion-based indoor localization. It uses both Wi-Fi and BLE fingerprints for sparsely calibrated location estimation. When applied to sparsely calibrated localization problems, our proposed method is advantageous in three respects. First, it dramatically reduces the human calibration effort required when using a semi-supervised learning framework. Second, it uses fused Wi-Fi and BLE fingerprints to markedly improve the location accuracy. Third, it inherits the beneficial properties of ELMs in terms of training and testing speed because the input weights and biases of hidden nodes can be generated randomly. The two main contributions of this paper can be summarized as follows: (1) The findings indicate that effective multi-data fusion can be achieved not only through data-layer fusion, feature-layer fusion and decision-layer fusion but also through the fusion of constraints within a model. (2) For semi-supervised learning problems, it is necessary to combine the advantages of different types of data by optimizing the model's parameters. These two contributions will be valuable for solving other similar problems in the future.

In our future work, we plan to apply this method in a more complex environment and compare it against more up-to-date algorithms to test its performance and robustness. Furthermore, the question of how to optimize the weighting coefficients during the training phase remains an open issue. We plan to find a heuristic method for this purpose. Subsequently, we will test FSELM on other signal-based indoor localization problems.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China under Grant Nos. 61572471, 61472399 and 61572004 and by the Science and Technology Planning Project of Guangdong Province, China, under Grant No. 2015B010105001.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflicts of interest regarding the publication of this paper.

Ethical approval This article does not report on any studies with human or animal participants performed by any of the authors.

References

Aparicio S, Pérez J, Bernardos AM, Casar JR (2008) A fusion method based on bluetooth and wlan technologies for indoor location. In: Multisensor fusion and integration for intelligent systems, 2008. MFI 2008. IEEE international conference on. IEEE, pp 487–491

Aparicio S, Pérez J, Tarrío P, Bernardos A, Casar J (2009) An indoor location method based on a fusion map using Bluetooth and WLAN technologies. In: International symposium on distributed computing and artificial intelligence 2008 (DCAI 2008). Springer, pp 702–710

Bahl P, Padmanabhan VN (2000) Radar: an in-building RF-based user location and tracking system. In: INFOCOM 2000. Nineteenth annual joint conference of the IEEE computer and communications societies. Proceedings. IEEE, vol 2, pp 775–784

Belkin M, Niyogi P, Sindhwani V (2006) Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. J Mach Learn Res 7:2399–2434

Chai X, Yang Q (2005) Reducing the calibration effort for location estimation using unlabeled samples. In: Pervasive computing and communications, 2005. PerCom 2005. Third IEEE international conference on. IEEE, pp 95–104

Chai X, Yang Q (2007) Reducing the calibration effort for probabilistic indoor location estimation. IEEE Trans Mob Comput 6(6):649–662

Chen J, Wang C, Sun Y, Shen XS (2011) Semi-supervised laplacian regularized least squares algorithm for localization in wireless sensor networks. Comput Netw 55(10):2481–2491

Chen Y-C, Chiang J-R, Chu H, Huang P, Tsui AW (2005) Sensor-assisted Wi-Fi indoor location system for adapting to environmental dynamics. In: Proceedings of the 8th ACM international symposium on modeling, analysis and simulation of wireless and mobile systems. ACM, pp 118–125

Chen Y, Yang Q, Yin J, Chai X (2006) Power-efficient access-point selection for indoor location estimation. IEEE Trans Knowl Data Eng 18(7):877–888

Chen Z, Chen Y, Gao X, Wang S, Hu L, Yan CC, Lane ND, Miao C (2015) Unobtrusive sensing incremental social contexts using fuzzy class incremental learning. In: Data mining (ICDM), 2015 IEEE international conference on. IEEE, pp 71–80

Chung FRK (1997) Spectral graph theory, vol 92. American Mathematical Society, Providence

Cooper M, Biehl J, Filby G, Kratz S (2016) Loco: boosting for indoor location classification combining Wi-Fi and BLE. Pers Ubiquitous Comput 20(1):83–96

Ding S, Zhang N, Zhang J, Xu X, Shi Zhongzhi (2017) Unsupervised extreme learning machine with representational features. Int J Mach Learn Cybern 8(2):587–595

Galván-Tejada CE, Carrasco-Jiménez JC, Brena RF (2013) Bluetooth–WiFi based combined positioning algorithm, implementation and experimental evaluation. Procedia Technol 7:37–45

Galvan-Tejada I, Sandoval EI, Brena R et al (2012) Wifi bluetooth based combined positioning algorithm. Procedia Eng 35:101–108

Gao X, Hoi SCH, Zhang Y, Wan J, Li J (2014) Soml: sparse online metric learning with application to image retrieval. In: AAAI, pp 1206–1212

González E, Prados L, Rubio A, Segura J, de la Torre Á, Moya J, Rodríguez P, Martín J (2009) Atlintida: a robust indoor ultrasound location system: design and evaluation. In: 3rd symposium of ubiquitous computing and ambient intelligence 2008. Springer, pp 180–190

Gu B, Sheng VS (2017) A robust regularization path algorithm for v-support vector classification. IEEE Trans Neural Netw Learn Syst 28(5):1241–1248

Gu B, Sheng VS, Li S (2015) Bi-parameter space partition for cost-sensitive svm. In: IJCAI, pp 3532–3539

Gu B, Sheng VS, Tay KY, Romano W, Li Shuo (2015) Incremental support vector learning for ordinal regression. IEEE Trans Neural Netw Learn Syst 26(7):1403–1416

Gu B, Sheng VS, Wang Z, Ho D, Osman S, Li S (2015) Incremental learning for ν-support vector regression. Neural Netw 67:140–150

Gu Y, Chen Y, Liu J, Jiang X (2015) Semi-supervised deep extreme learning machine for Wi-Fi based localization. Neurocomputing 166:282–293

Haeberlen A, Flannery E, Ladd AM, Rudys A, Wallach DS, Kavraki LE (2004) Practical robust localization over large-scale 802.11 wireless networks. In: Proceedings of the 10th annual international conference on mobile computing and networking. ACM, pp 70–84

Ham J, Lee DD, Saul LK (2005) Semisupervised alignment of manifolds. In: AISTATS, pp 120–127

Han D, Hu Y, Ai S, Wang G (2015) Uncertain graph classification based on extreme learning machine. Cogn Comput 7(3):346–358

Hossain AKMM, Van HN, Jin Y, Soh W-S (2007) Indoor localization using multiple wireless technologies. In: Mobile adhoc and sensor systems, 2007. MASS 2007. IEEE international conference on. IEEE, pp 1–8

Huang G-B, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern Part B (Cybern) 42(2):513–529

Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1):489–501

Lee S, Ha KN, Lee KC (2006) A pyroelectric infrared sensor-based indoor location-aware system for the smart home. IEEE Trans Consum Electron 52(4):1311–1317

Letchner J, Fox D, LaMarca A (2005) Large-scale localization from wireless signal strength. In: AAAI, pp 15–20

Lin Q, Zhao F, Luo H, Kang Y (2011) A wireless localization algorithm based on spectral decomposition of the graph laplacian. Acta Autom Sin 37(3):316–321

Liu H, Darabi H, Banerjee P, Liu J (2007) Survey of wireless indoor positioning techniques and systems. IEEE Trans Syst Man Cybern Part C (Appl Rev) 37(6):1067–1080

Liu J, Chen Y, Liu M, Zhao Z (2011) Selm: semi-supervised elm with application in sparse calibrated location estimation. Neurocomputing 74(16):2566–2572

Liu M, Liu B, Zhang C, Wang W, Sun W (2017) Semi-supervised low rank kernel learning algorithm via extreme learning machine. Int J Mach Learn Cybern 8(3):1039–1052

Liu Y, Zhang L, Deng P, He Z (2017) Common subspace learning via cross-domain extreme learning machine. Cogn Comput 9(4):555–563

Lowe D (1988) Multi-variable functional interpolation and adaptive networks. Complex Syst 2:321–355

Mao W, Wang J, Xue Z (2017) An elm-based model with sparse-weighting strategy for sequential data imbalance problem. Int J Mach Learn Cybern 8(4):1333–1345

Nguyen X, Jordan MI, Sinopoli B (2005) A kernel-based learning approach to ad hoc sensor network localization. ACM Trans Sens Netw (TOSN) 1(1):134–152

Ouyang RW, Wong AK-S, Lea C-T, Chiang M (2012) Indoor location estimation with reduced calibration exploiting unlabeled data via hybrid generative/discriminative learning. IEEE Trans Mob Comput 11(11):1613–1626

Pan JJ, Pan SJ, Yin J, Ni LM, Yang Qiang (2012) Tracking mobile users in wireless networks via semi-supervised colocalization. IEEE Trans Pattern Anal Mach Intell 34(3):587–600

Pan JJ, Yang Q, Chang H, Yeung D-Y (2006) A manifold regularization approach to calibration reduction for sensor-network based tracking. In: AAAI, pp 988–993

Pan JJ, Yang Q, Pan SJ (2007) Online co-localization in indoor wireless networks by dimension reduction. In: Proceedings of the national conference on artificial intelligence. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, vol 22, p 1102

Pandya D, Jain R, Lupu E (2003) Indoor location estimation using multiple wireless technologies. In: Personal, indoor and mobile radio communications, 2003. PIMRC 2003. 14th IEEE proceedings on. IEEE, vol 3, pp 2208–2212

Pao Y-H, Park G-H, Sobajic DJ (1994) Learning and generalization characteristics of the random vector functional-link net. Neurocomputing 6(2):163–180

Rao CR, Mitra SK (1972) Generalized inverse of matrices and its applications. Wiley, New York

Rodrigues ML, Vieira LFM, Campos MFM (2012) Mobile robot localization in indoor environments using multiple wireless technologies. In: Robotics symposium and Latin American robotics symposium (SBR-LARS), 2012 Brazilian. IEEE, pp 79–84

Scardapane S, Comminiello D, Scarpiniti M, Uncini A (2016) A semi-supervised random vector functional-link network based on the transductive framework. Inf Sci 364:156–166

Schmidt WF, Kraaijveld MA, Duin RPW (1992) Feedforward neural networks with random weights. In: Pattern recognition, 1992. Conference B: pattern recognition methodology and systems, proceedings, 11th IAPR international conference on. IEEE, vol II, pp 1–4

Serre D (2002) Matrices: theory and applications. In: Graduate texts in mathematics. Springer, New York

Torres-Solis J, Falk TH, Chau T (2010) A review of indoor localization technologies: towards navigational assistance for topographical disorientation. INTECH Open Access Publisher

Vapnik V (2013) The nature of statistical learning theory. Springer science & business media, Berlin

Wen X, Shao L, Xue Y, Fang W (2015) A rapid learning algorithm for vehicle classification. Inf Sci 295:395–406

Wong PK, Gao XH, Wong KI, Vong CM (2016) An analytical study on reasoning of extreme learning machine for classification from its inductive bias. Cogn Comput 8(4):746–756

Xiang L, Wang D, Wei Y, Zhou Y (2015) Location-fingerprint based indoor localization via scalable semi-supervised learning. Int Inf Inst (Tokyo) Inf 18(2):641

Xu L, Ding S, Xu X, Zhang N (2016) Self-adaptive extreme learning machine optimized by rough set theory and affinity propagation clustering. Cogn Comput 8(4):720–728

Yang Z, Wu C, Liu Y (2012) Locating in fingerprint space: wireless indoor localization with little human intervention. In: Proceedings of the 18th annual international conference on mobile computing and networking. ACM, pp 269–280

Zhai J, Zhang S, Wang C (2017) The classification of imbalanced large data sets based on mapreduce and ensemble of elm classifiers. Int J Mach Learn Cybern 8(3):1009–1017

Zhang L, Deng P (2017) Abnormal odor detection in electronic nose via self-expression inspired extreme learning machine. IEEE Trans Syst Man Cybern Syst PP(99):1–11

Zhang L, He Z, Liu Y (2017) Deep object recognition across domains based on adaptive extreme learning machine. Neurocomputing 239:194–203

Zhang L, Zhang D (2015) Domain adaptation extreme learning machines for drift compensation in e-nose systems. IEEE Trans Instrum Meas 64(7):1790–1801

Zhang L, Zhang D (2016) Robust visual knowledge transfer via extreme learning machine-based domain adaptation. IEEE Trans Image Process 25(10):4959–4973

Zhang L, Zhang D (2017) Evolutionary cost-sensitive extreme learning machine. IEEE Trans Neural Netw Learn Syst 28(12):3045–3060

Zhang Y, Zhi X (2010) Indoor positioning algorithm based on semi-supervised learning. Comput Eng 36(17):277–279

Zhou M, Tian Z, Xu K, Yu X, Hong X, Wu H (2014) Scanme: location tracking system in large-scale campus Wi-Fi environment using unlabeled mobility map. Expert Syst Appl 41(7):3429–3443